

Game Theory in Christian Perspective: The Cooper Project?

Geoffrey Brennan and Christopher White
Australian National University

In a recent issue of this journal, Ben Cooper (2015) has laid out an intelligent and wide-ranging account of “game theory” as a potentially useful explanatory and normative device in Christian social and economic thinking. The aim of our commentary is to adumbrate and clarify that account, rather than to criticize—though there are some aspects of Cooper’s exposition that we think are misleading. We will begin our discussion with one example of this latter kind. We will then go on to underline a number of distinctions we think it useful to maintain in the interpretation of game theory. Some of these track Cooper’s treatment quite closely. Others do not.

1 Interdependence of What?

At the outset, Cooper (2015, 1) describes game theory as

... a discipline which considers the behavior of decision-makers in interactive situations, focusing on cases where there is more than one decision-maker, where the outcome depends upon the combination of decisions, and where each decision involves considering what the other decision-makers will choose.

It would be possible to quibble here over whether game theory is itself a “discipline” or is better thought of as an “approach” or an “analytic technique” that is deployed in a variety of domains more commonly thought of as disciplines (e.g. economics, political science, sociology, social psychology, and evolutionary biology). Exactly how the category of “discipline” is to be understood and what marks off one discipline from another within the organization of enquiry is an interesting question; but it is not one that will occupy us here—beyond recording a nagging suspicion that game theory might not qualify¹.

In the quotation given above, Cooper lists a trio of properties that to-

gether define game theory. Our substantive quibble focuses on the third element—what we might refer to as the distinctively “strategic” element in the cases with which game theory deals. This is the idea that in choosing how to act, the rational individual is required to make predictions about the behaviour of other actor(s) in the game. Cooper’s definition here tracks, to be sure, what many game theorists *say* when they describe their enterprise, but this definition diverges rather spectacularly from what they actually *do*. In particular, some of the most common forms of “interactive situations” with which game theory deals—including in particular the prisoner’s dilemma²—do *not* involve “*each decision[-maker] ... considering what other decision-makers will choose.*” These are games in which at least some decision-makers have a “dominant strategy”—an action choice that is best for the chooser whatever the other chooses. Since such ‘games’ are taken to be representative of many important social situations (and are explicitly discussed by Cooper in the substance of his paper), it seems bizarre to adopt a definition of game theory that effectively excludes them by stipulation at the outset.

To be fair, Cooper’s definition does not specify that actors are “*rationaly required*” to consider what other decision-makers will choose, but merely that they *do* so consider. But in games where both players have a dominant strategy (like the prisoner’s dilemma) any such “considering” will be entirely gratuitous. The critical point is that, in this class of (genuine) games, decision-makers do not have any reason to consider how others will choose.

In any event, Cooper is later quite explicit on the relevant point. On page 2 he reiterates the relevant aspect of his definition. He points out that:

For a given agent (now called a “player”), the resulting outcome depends not just on that player’s choice, but on the choices of the other player or players.

This much is entirely correct. But Cooper immediately goes on to say:

The choice a player makes is therefore influenced by what he or she believes the other players will choose.

And this second “*therefore*” claim simply does not follow—as the standard prisoner’s dilemma illustrates. That claim is a logical error.

The general point is this: Game theory deals with the structure of in-

ter-dependence between decision-makers. So much is unexceptionable. But the interdependence can be in terms of *action* or in terms of *payoff* structure. In the n-person prisoner's dilemma, for example, the rational action of each is independent of the action of others, but the payoff of each is *not* independent.³ The payoff to each depends on what others do, and that is so for all players. So even though the behavioural choices that each makes do not require asking what others will do, each is nevertheless *affected* by what others do. When, for example, you emit carbon, the rest of the world suffers from the atmospheric degradation thereby created, and you certainly endure the atmospheric degradation caused by the emission actions of others (as well as your own). But the net "payoff" to each from her own reduction in carbon emissions (if any) is such that it does not pay any of us to reduce emissions unilaterally: Each does not need to know what others are doing or how others might respond to her actions in order to know that it is better for her material well-being to emit than to refrain from emitting. There is no *action*-interdependence; but there is *payoff*-interdependence.

Why might this distinction between action- and payoff-interdependence matter? Because it connects to a distinction between three different purposes to which game theory might be put:

1. The purpose might be to explain (or predict) behavior within the context of the interaction;
2. The purpose might be to explain (or predict) behaviour at the more abstract level of responding to the fact of the predicament itself;
3. The purpose might be essentially normative—to analyse what people *ought* to do, either within the game or in response to the predicament that the game reveals.

To exemplify, consider again the prisoner's dilemma case, as illustrated by carbon emissions. You might be interested to predict behavior in relation to carbon emissions when the costs of any increase in your emissions are almost entirely borne by others. (Purpose 1) That could be a rather uninteresting (and disappointing) exercise if everyone is motivated exclusively by payoff maximization (as is often assumed), since it follows from the claim that the carbon emissions case has the prisoner's dilemma payoff structure that everyone will emit to the point that is in their individual interest—and well beyond the level that is collectively optimal.

(We shall have more to say about this motivational issue in Section 3.) Or you might want to explain why, given the prisoner's dilemma structure of the game, countries might have reason to make agreements to reduce their emissions and try to enforce those agreements both on other countries and on their own citizens. (This is purpose 2.)⁴ Or you might want to work out what is the right thing to do in the face of the prisoner's dilemma predicament. And that "right thing" might include undertaking morally responsible action yourself, encouraging others to do so as well, encouraging your government to introduce carbon emission restrictions (perhaps via a carbon tax), and / or encouraging your government to be pro-active in pursuing international agreements about introducing such policies across the globe.

If your purpose were the first, then you might wish to focus attention on cases where what each chooses to do depends on what others choose to do. So you might want to restrict game theory to those cases where the *actions* of players are interdependent. That is what Cooper's specification strictly interpreted involves.

However, we think that Christians in general would not wish to rule out the normative purposes of game theory. Most Christians want to understand the way the social order works in order to make that order work better. So, in fact, do many economists—and others who use game theory as a tool. Christians may, of course, have a distinctive understanding of what "better" means. We shall need to say a little about that issue—the Christian understanding of "betterness"—in Section 2, because it is relevant to how Christians might use game theory. But the basic point here is that securing an outcome in prisoner's dilemma predicaments that generates the higher payoff is *presumptively* a matter of normative significance. Appealing again to the carbon emissions case, securing a situation in which carbon concentrations in the atmosphere are reduced to levels consistent with acceptable climatic conditions is something that everyone has reason to want to achieve—Christian and non-Christian alike. The grounds for that outcome offered by Christians might differ from those offered by a secular utilitarian. But here, as in many other cases, Christians have common cause with all people of good will. Achieving outcomes that are preferred by everyone is something that is ethically desirable in many cases, on any plausible rendering of "ethically desirable".

2 Payoff Maximization and Ethics

We have said that, in order to incorporate the full range of normative concerns that game theory is useful in illustrating, game theory ought to be defined in terms of payoff interdependence, rather than action interdependence⁵. Specifically, game theory ought to be defined to allow for prisoner's dilemma interactions,⁶ where there is no action interdependence in the terms that Cooper's definition specifies. But it is worth noting that the claim that prisoner's dilemma predicaments are helpful in illuminating normatively relevant social situations commits one to the view that, in many cases, increasing payoffs to the players is ethically desirable. That is, the payoffs themselves are of normative significance.

But that is not always so, and it is important to show this because we might otherwise import normative presumptions where none is justified. Cooper himself offers an example, but he uses it to draw a tendentious conclusion.

Cooper's example is this: A drug lord and a cartel of local suppliers are engaged in a potential transaction involving the purchase of drugs. But in this exchange there is an incentive on both sides to defect. The suppliers can adulterate the drugs; the drug lord can pay in counterfeit money that will turn out to be worthless. Both parties know of the risk of defection and so the transaction doesn't occur, even though both parties would be better off if the unadulterated drugs were exchanged for real cash. (Call this latter case "the 'deal.'") The interaction has the structure of a prisoner's dilemma. And on that basis, Cooper seems to think that a simple utilitarian ethicist (or the standard economist) would be obliged to conclude that the deal should go through. A Christian by contrast would not be so obliged, because the Christian is not committed to payoff maximization as the only normatively relevant consideration.

However, this is only half right. It is true that the Christian's normative framework is concerned with more than payoff maximization; but it is not true that the utilitarian (or the economist for that matter) would be obliged to think that the failure of the deal is a bad thing. Utilitarians are not so obliged because the drug transaction affects more people than those involved in the deal. At the very least, it affects the drug addicts who would be the final users. And it affects the drug addicts' dependants. Such effects are normatively relevant and they imply, as a matter of simple logic, that identifying a situation as a prisoner's dilemma *for a certain*

set of actors is not necessarily the end of normative evaluation (even for, and perhaps especially for, a utilitarian!).

Indeed, this is a point that economists are familiar with, at least in one specific case. Perfect competition, which economists generally think of as a desirable institutional form for many ordinary market transactions, depends on the instability of producer cartels. It depends, that is, on the relation between members of a potential cartel having a prisoner's dilemma structure, so that it will be in the interests of each member to charge a lower price than the price that maximizes profits across cartel members. (By doing so, any firm can expect to gain at the expense of all those who maintain the monopoly price). This case, at least, provides one important (and familiar) instance where the prisoner's dilemma is good, not bad. And it is so, because the parties to the cartel are not the only individuals who matter normatively.

The bottom line here is that there are two reasons why prisoner's dilemmas may *not* be bad:

- a) Because the "players" in the prisoner's dilemma are not the only parties affected by the outcome of the interaction. Utilitarians, economists, and Christians alike would want to include the effects on all those affected. Focussing on the payoffs to *players* alone in any case where there are others affected is too narrow.
- b) Because the understanding of ethical desirability deployed by the Christian is rather different from that deployed by utilitarians and typically by economists. As Cooper rightly points out, economists are much disposed to take individuals' preferences at face value. So when the prisoner's dilemma produces outcomes that the players do not prefer (as it characteristically does), economists are inclined to think that that's bad news *for the players*.⁷ Christians are certainly not committed to that view. Christians are, we take it, committed to a more substantive view of what constitutes human flourishing—one in which certain attitudes and beliefs and intentions play a central role.

This is hardly the occasion on which to develop even a minimal account of the content of Christian ethics and how it differs from its secular rivals. What we would emphasize here is just the following point: The fact

that Christian ethics considers more than payoffs (and / or preference satisfaction) does not imply that payoffs or preference satisfaction are normatively irrelevant. That man does not live by bread *alone*, as Jesus reminds Satan, does not deny that bread is good—something which the various accounts of Jesus’ miracles involving loaves and fishes suggest Jesus himself believed.⁸

This point relates to another example that Cooper offers, which we think deserves further development. Cooper considers the case in which individuals are motivated by a desire for “social status”.⁹ Cooper’s thought is that both players can be better off if they agree to moderate the “rat-race” that each is involved in, because when both strive equally hard to outdo the other, they end up, in *relative* terms, in the same relation as they would have been in if they had both striven proportionately less. So, Cooper concludes, buying “useful” goods rather than “status” goods will be better for both. The thought here is that by changing individuals’ preferences—each becoming less preoccupied with how she stands in the eyes of others—the prisoner’s dilemma structure can be subverted: each can achieve a larger material payoff than she otherwise would.

We want to put pressure both on this particular example and the general argumentative structure that it represents. So consider the example first. For the purposes of the exercise, let us concede that social status is itself a positional good (either directly in that status reflects an ordinal ranking or indirectly in that social approval is fixed in aggregate terms, so that an increase in your level of approval will reduce others’ social approval, other things equal). After all, social status does reflect the approval that third parties provide when a player acts in socially approved ways. And that social approval may be more or less justified. If observers themselves have debased values (so, for example, they give A approval based on the age and luxury of A’s car) then social approval will reflect that fact. But suppose social approval is given for honesty or kindness or professional conscientiousness (as it surely normally is in all three cases). Then acting to achieve that social approval will be acting in ways that are properly normatively approvable. And if the desire for social approval stimulates both A and B to invest greater energy in being honest / kind / professionally proper, then the net effects of the desire for social approval / status are likely to be positive.¹⁰ Here, we are making appeal to point a) above. The mere observance of a prisoner’s dilemma structure for the players (the status-seekers) does not settle

the issue as to whether status-seeking produces desirable consequences or not.¹¹

It is, however, the argumentative structure that is our chief source of anxiety here. The point of Cooper's treatment seems to be to suggest that changes in agent motivation of the kind seen by St. Paul as characteristic of "Christian virtue" have the effect of solving prisoner's dilemma predicaments—and perhaps by implication other kinds of game-theoretic problems as well. But even supposing that that claim were true (which we do not concede), what is the force of any such observation? That this fact gives extra reason to adopt the Christian virtues? That it provides evidence that Paul was right in specifying Christian virtue in the way he does?

The first possibility involves a logical error related to the prisoner's dilemma itself, because what maximizes my payoff is that *everyone else* be virtuous. If I'm the only virtuous person among the group, the prisoner's dilemma isn't "solved"; it just ensures that I am exploited. So each individual might wish that others become virtuous, but each has no reason derived from prisoner's dilemma logic to become virtuous herself.

And the second possibility seems just wrong-headed in itself. It seems like an attempt to reduce Christian ethics to a form of rule-consequentialism. It makes the final arbiter of different motivational structures whether they solve various kinds of dilemmas or not—that is, whether they allow players to achieve outcomes that involve maximal mutual-objective payoffs. But what doth it profit a man to gain a maximal objective payoff (or a minimal prison sentence) and lose his soul? If that question is taken seriously, then it remains unclear just what the examples Cooper deploys are supposed to show.

Just for the record in this connection, it is worth pointing out that there is an important category of games in which changes in motivation—at least of the kind St. Paul is discussing—have no effect. This is the range of coordination games, the characteristic feature of which is that each player does need to predict what the other is going to do in order to determine what is best for her to do. The classic example is deciding on which side of the road to drive, in some context where there is not already an established convention. In such settings no amount of altruism or benevolence alters the basic structure. You may desire that coordination be secured because it is good for the other (as well as or rather than for yourself). But that feature is a fifth wheel: it does not alter the nature of the predicament and it does not in any way proffer a solution!

But the more significant problem in Cooper's approach lies not so much in cases where Christian virtue cannot solve predicaments that game theory isolates, but where changes in motivation *can* solve predicaments, and where that fact is used as a mechanism for inducing the content of Christian virtue. The problem is worth an extended illustration.

Consider the original classic version of the prisoner's dilemma in which each of two (alleged) criminals has a choice of whether to confess or not confess to a particular crime. The district attorney faces each with the following payoff structure:

If you confess and your partner doesn't, he'll get ten years and
you'll go free;
If you both confess, you'll each get eight years;
If neither confesses, you'll each get one year for a lesser crime.

There is a dominant strategy for each player: namely, to confess. Both confessing is the Nash equilibrium in this game and is clearly worse for each player than if both kept quiet.

Suppose now we consider the operatic version of this game, in which the lovers Cavaradossi and Tosca are arraigned before the wicked Scarpia. However, being opera, Cavaradossi cares only for Tosca and Tosca cares only for Cavaradossi. Each seeks to minimize the sentence for the other. Then, under the original payoff structure, neither confesses. In that way the beloved gets at most one year. But Scarpia, being smart as well as evil, and knowing that this is the way the lovers will choose, simply reverses the sentences:

If you confess and your lover doesn't, you'll get ten years but s/he
will go free;
If both confess, eight years each;
If neither confesses, one year each as before.

Subject to both behaving "rationally" as defined, both will confess, preserving the characteristic feature that the Nash equilibrium is 'Pareto-dominated' (that is, the game gives rise to a less-desired outcome for both players than might have been achieved if both had stayed silent). Perhaps the example is a little far-fetched; but the general point is a logical one, not an empirical one. It is, after all, a common thought that attending only to the interests of others is more virtuous a motivation than

attending exclusively to one's own interests. One might think that such extravagant self-sacrificial love is precisely what Christians are called to (as Cooper perhaps implies in his reference to John 13:34–35). Perhaps that is so, perhaps not. Our point is simply that such virtue does not make a community impervious to prisoner's dilemma predicaments.

A possible response is that such problems would not arise for any player who treated her neighbour *as herself*. Such a player would not disregard the interest of the other, but neither would she disregard her own interest (as Paul says in Philippians 2:4). She might, for example, simply add the payoffs for each possible outcome and choose the action that had the best expected total payoff.¹² Sometimes this motivation will produce “cooperation” in prisoner's dilemma interactions: sometimes it will confront each player with a coordination game (see below).¹³ Perhaps one might appeal to such considerations as giving reason for “treating neighbor and self symmetrically” rather than “counting oneself as of nothing” as in the Cavaradossi / Tosca case.

But we think that this is a very slim reed indeed on which to base the correct interpretation of Christian virtue! And if that is so, one must wonder exactly what to make of Cooper's argumentative strategy. To be sure, sometimes players who exhibit Christian virtue will be able to avoid prisoner's dilemma predicaments. But sometimes they won't. And sometimes what might plausibly be thought of as Christian virtue will render players subject to prisoner's dilemma predicaments that other interpretations of Christian virtue might avoid.

So what? The most natural interpretation of Cooper's argumentative strategy involves the thought that game theory can be deployed to show why Christian virtue is a “good thing.” But that strategy serves to elevate “payoff maximization” to the status of a trump card (a sort of ultimate justifier) in Christian interpretation. And that is just a mistake—and a serious one, in our view.

3 Preferences vs. Payoffs

Cooper draws a distinction (one we think important) between preference satisfaction and dollar payoff maximization, especially in relation to the ultimatum game. He thinks (totally correctly in our view) that the assumption that players will maximize the monetary value of the payoffs is a substantive assumption about the content of preferences; it is not

something that follows from the rationality of players. The proper conclusion to be drawn from the ultimatum game, for example, is not that players are typically “irrational” but rather that there is more at stake for the players than monetary payoffs.

The manner in which Cooper prefers to deploy terms is to define “rationality” as requiring “payoff” maximization,¹⁴ but to insist that payoffs must reflect the full benefit in terms of preference satisfaction that is associated with different actions (by all players). So, in laboratory experiments of the ultimatum game, for example, most second-movers veto offers that give them anything less than a third of the total stake. Cooper’s formulation involves insisting that the players “*are playing a different game from the ultimatum game*” (2015, 11).

Our own terminological preference here is to allow the game to specify the payoff structure in monetary or equivalent terms,¹⁵ but to leave it an open question as to how players will behave in the face of the monetary (or other material) incentives that different strategies offer. The payoffs specified within the game might be expressed in terms of dollars or (as in the classic prisoner’s dilemma narrative) years in prison or (as in the evolutionary case) survival value or probability of reproduction. But the structure of what we might call these “objective” payoffs can be used to define the game, leaving it open as to whether behavior will track those objective payoffs (or whether alternatively there might be “subjective” factors that will lead to outcomes other than those the “objective payoff maximization” hypothesis would predict).

Our reason for this preference is twofold. First, we think it is useful to maintain a distinction between objective payoffs and subjective ones for *normative* purposes. Typically, the normative agenda in game theory is to reconstruct material incentives so as to secure a specific outcome where the objective payoffs have appropriate normative features (maximal aggregate material payoffs or desirable distributive properties or some such). The exercise is to devise material incentives to achieve those goals, as with taxes on carbon emissions or subsidies for carbon substitutes or penalties for littering.

Second, the social predicament that the game is designed to illustrate is often characterized by the interdependent structure of those *objective* payoffs, and something important is lost if we do not allow that structure to define the interaction. So, for example, in the climate change predicament, we think it important to represent the situation as an n-person

prisoner's dilemma, even if many individuals of good conscience make some efforts to reduce their own individual emissions. For example, representation in terms of material payoffs indicates why the same prisoner's dilemma structure applies not just to each individual carbon-emitting individual but also to nation-states. We are thereby alerted to the following challenge: If individuals "free-ride" in (not) reducing their emissions, then we have some reason for thinking that national governments¹⁶ will also do so, and for exactly the same reasons. And so we are led to an understanding as to why international treaties to reduce carbon emissions are so difficult to consummate (or if treaties are made, why countries might prove recalcitrant in sticking to their terms).

It is surely the case some individuals will choose to reduce their own emissions (say by consuming non-carbon-intensive products or by buying carbon off-sets¹⁷) even though it is not in their narrow interests to do so. Such individuals may be motivated by Christian principles or moral considerations of other kinds; they may want to have the satisfaction of "keeping their own hands clean." Larger numbers of individuals may express a similar preference for atmospheric hygiene at the ballot-box, placing some pressure on their governments to introduce policies designed to effect carbon emission restrictions on a national scale. But so far, the trend in global carbon emissions (and corresponding atmospheric carbon concentrations) has not reversed: sadly, individual carbon reducers remain decidedly in the minority. The incentives that are characteristic of the n-person prisoner's dilemma explain why this outcome is likely to arise, and also suggest why multi-national agreements might be required in order to secure a genuine solution to the world's looming climatic problems.¹⁸

So, like Cooper, we want to distinguish between the objective payoff structure specified in a game and the subjective preferences that determine how individuals will act. And in the same spirit, we think it a straight-forward mistake to identify "rationality" with "objective payoff" maximization. But since the role of game theory in many significant social applications is to diagnose predicaments of one kind or another, and since the objective payoffs often play an important role in defining those predicaments, we think those objective payoffs ought to be retained (alongside subjective ones). Of course, this usage betrays an underlying assumption that objective payoffs play *some* role (and sometimes a predominant one) in motivating players' behavior in many social interactions.¹⁹

But like Cooper (and presumably most Christians and many non-Christian moralists) we do not think that individuals are *exclusively* motivated by self-interest, and it is just a descriptive mistake (and sometimes a very major one) to make that assumption. As Alexander Hamilton puts the point in Federalist 76:

The supposition of universal venality is little less an error in human reasoning than the supposition of universal rectitude.

Or as the Christian might put the point, both sin and virtue belong in any proper Christian anthropology. To ignore either is to overlook something of great significance, descriptively as well as theologically.²⁰

4 Normative and Theological Considerations

Christians are most likely to be interested in game theory because game theory is useful in exposing features of social interactions that are normatively significant. Of course, Christians have no monopoly on such normative concern. Virtually all public policy is driven by normative considerations, and game theory helps reveal the structural nature of many important public policy challenges. However, as Cooper emphasizes, the normative approach used by economists in public policy tends to be broadly utilitarian, and certainly consequentialist. And he is perfectly correct to register that “ethical assessment from a Christian point of view” will be concerned *as well* with the “virtue” of agents—their motives and modes of calculation, and their attitudes. At the same time, we think he is also right to emphasize that a concern with Christian virtue does not obliterate a concern with outcomes. As Jesus emphasizes, the Sabbath was made for man, not man for the Sabbath! So, the general conclusion is that, for the Christian, outcomes matter, and material well-being matters, but these are not the *only* things that matter. Christian normative evaluation differs importantly from more typical economic normative frameworks in insisting that that is so.

In what we take to be the most plausible rendering of a Christian normative position, there are two elements that are of significance:

1. Individual moral attitudes and conduct—what we might think of as individual virtue (understood in distinctively Christian terms). When we refer to “attitudes” here, we have in mind such things as an awareness of one’s own creatureliness and

sinful nature and need of redemption, as well as a conviction that such redemption becomes possible through the death and resurrection of Christ.

2. The properties of the social outcomes to which individual interactions give rise. These properties will include the material wellbeing of the human race (and especially of those members whose levels of well-being are lowest), with 'well-being' here interpreted broadly to include health and liberty, as well as economic well-being and fairness in the distribution of these things.

These two broad elements interconnect. Sometimes they will be complementary: efforts to make progress on the first will have positive effects on the second. Sometimes they will be essentially independent, as we claim is the typical case in coordination predicaments. And sometimes, as Mandeville (1714) notoriously contended, they will be in conflict. And even where they are not in direct conflict, choices will have to be made as to where to locate one's efforts. Often "compromises" will be required, and those compromises will often involve ways of securing as much of the second as can be achieved without making heavy demands on the first. (For example, if climate change / carbon emission problems cannot be solved by appeal to Christian virtue alone, then Christians might plausibly seek to find ways of solving those problems in ways that make minimal demands on virtue. This enterprise is what one might characterize as "economizing on virtue," and on certain readings it is largely in this that good policy design in the economic tradition resides.)

There is a danger here to which we need to be alert. If one thought that it was axiomatic that Christian virtue promotes best outcomes, then it would be possible to work backwards from those motivations that are maximally effective in solving various social problems to an understanding of what true Christian virtue consists in. This seems to be the object of one strand of Cooper's discussion. Indeed, for him, throwing light on what St. Paul has to say about Christian virtue seems to be one of the central points at which game theory and theology come together. We think that latter project is misconceived.²¹ We think that proper understanding of Christian virtue is an independent exercise, in which a good understanding of game theory plays at best a minimal role.

5 The Bottom Line

Game theory can be useful in understanding structures of interdependence and in isolating cases where material incentives are ill-tuned to achieving good outcomes. As we see it, the Christian normative scheme recognizes a proper place for such understanding. Presumably, for some Christians, a proper conception of their Christian vocation might include developing ways in which better outcomes can be achieved in the mixed world of human sin and human virtue that we inhabit. And Christians can perhaps contribute something distinctive to the study of game theory by being more sensitive to the fact that players are often motivated by a wider range of concerns than their own material payoffs. But the idea that game theory can illuminate the idea of Christian virtue by focussing on cases where putatively “virtuous” motivations avoid predicaments that would dog more selfish players—that idea is, we think, seriously misconceived.

Endnotes

- 1 To give a little substance to our misgivings here, we should declare a suspicion that enquiry within the social sciences is best driven by substantive concerns rather than by an analytic technique in search of applications.
- 2 In this commentary, we will refer to various “games,” such as the prisoner’s dilemma or the ultimatum game, without bothering to specify the structure that defines such games. Cooper has done this in ways that seem to us to be entirely sufficient.
- 3 Actually, in the prisoner’s dilemma case, a characteristic feature is the property that the payoff to each player depends more on what the other does than on what each herself does. It is perhaps for this reason that some commentators have thought there is a close resemblance between the prisoner’s dilemma and the situation involving mutually agreeable exchange in goods.
- 4 Note that this explanation depends on the idea that the carbon emission game *is* a prisoner’s dilemma (or something very like it). One wouldn’t need government action if it genuinely were in the interests of individual emitters to restrict their own emissions.
- 5 Any case of action interdependence will involve payoff interdepen-

- dence and so a definition of game theory in terms of payoff interdependence will be more inclusive.
- 6 And for that matter so-called “prisoner’s delight” interactions, which likewise involve dominant strategies for all players.
 - 7 The picture in relation to utilitarianism is more complicated. “Utility” may not be defined in terms of payoff maximization or preference satisfaction. It might instead be thought of as “well-being” — with the relation between an individual’s well-being and specific payoffs or preference satisfaction more generally left as an open question.
 - 8 Of course, those miracles are capable of a variety of interpretations, but we think all such interpretations are consistent with the simple proposition that, when the thousands are fed, Jesus has done something good!
 - 9 Cooper’s exposition tends to treat the desire for social status as co-terminous with envy. Those motivations are not strictly identical. And it is not clear that envy exhibits a prisoner’s dilemma structure: an outcome where both have a higher material payoff may not involve greater preference satisfaction to either, if each prefers to sacrifice some material benefit to impose a material loss on the other. So we shall focus on the social status case.
 - 10 The rationale of overt recognition in the form of awards for charitable and other public-spirited behaviour is just such a case in point.
 - 11 For a much more detailed treatment of related issues, see Brennan and Pettit *The Economy of Esteem* (2004). Adam Smith deals with a somewhat similar case to Cooper’s, in a famous passage in his *Theory of Moral Sentiments* [Part IV Chapter 1]. Smith refers to “*the poor man’s son, whom heaven in its anger has visited with ambition...*” who labours long and hard to achieve distinction, only to find at last “... *that wealth and greatness are mere trinkets of frivolous utility, no more adapted for procuring ease of body and tranquility of mind than the tweezer-cases of the lover of toys.*” But, Smith goes on to assert, “*it is well that nature imposes on us in this manner. It is this deception that rouses and keeps in continual motion the industry of mankind.*” The consequences of this “deception” are cities and commonwealths, arts and sciences, and an advancement in “*the interest of society and ... multiplication of the species.*”
 - 12 Which incidentally is exactly what a simple utilitarian would do—a

- fact worth noting, because utilitarianism has had at best an ambivalent reception in Christian circles!
- 13 Which it will be depends on the precise *values* of the various payoffs, not just the *ranking* of outcomes.
 - 14 “Rational” must be defined quite specifically in game theory, and generally does not have the same meaning as in common usage, such as “sensible.”
 - 15 In the classic prisoner’s dilemma, the payoffs are expressed as years in prison.
 - 16 It is perhaps not self-evident that it is not in the “national interest” of individual countries to introduce carbon emissions reductions unilaterally. That is something the prisoner’s dilemma representation makes more salient. We should perhaps not be surprised that the Copenhagen conference failed so dismally, or that the Kyoto round earlier produced largely token action by only a handful of governments.
 - 17 As John Broome recommends in *Climate Matters* (2014).
 - 18 Assuming that the scientific consensus on the relation between carbon concentrations and climate change are correct. Of course, part of the difficulty here is that by the time we know for sure that there is a problem it is likely to be too late to do much about it. What the prisoner’s dilemma representation shows us, however, is that even if everyone were certain as to the effects, each would be materially better off if the burden of carbon restriction fell upon everyone else. Of course, if the result is that everyone waits for someone else to act, in the longer run everyone is materially worse off.
 - 19 Simon Blackburn (1998, Chapter 6) draws a distinction between “empirical games” and “interpreted games” (or perhaps better between an empirical and an interpreted representation of the same game) in exactly the same spirit as our distinction between the objective and subjective representation.
 - 20 We do not mean to suggest that any action that increases one’s own material payoffs is presumptively sinful; merely that it forms part of any overall assessment of human virtue and sinfulness.
 - 21 It is not of, course, an entirely unfamiliar project in the history of moral philosophy: on one reading, that is precisely what rule utilitarianism attempts to do for utilitarianism.

References

- Blackburn, S.** (1998). *Ruling Passions*. Oxford: Oxford University Press.
- Brennan, G. and Pettit, P.** (2004). *The Economy of Esteem*. Oxford: Oxford University Press.
- Broome, J.** (2014). *Climate Matters*. Oxford: Oxford University Press.
- Cooper, B.** (2015). Game Theory in Christian Perspective. *Faith and Economics*, 65, 1–23.
- Manderville, B.** (1705/1957). *The Fable of the Bees*. Oxford: Clarendon Press.
- Smith, A.** (1759/1976). *Theory of Moral Sentiments*. Oxford: Oxford University Press. ■