Cooperative Games and Cooperative Organizations

By Roger A. McCain
Drexel University, Philadelphia, PA
mccainra@drexel.edu

Abstract

It is well known that game theory has two major branches, cooperative and noncooperative game theory. Noncooperative game theory is the better known and more influential of the two. A difference is that cooperative game theory admits of binding agreements to choose a joint strategy in the mutual interest of those who agree. Cooperative organizations, too, are seen as being in the mutual interest of the members, but there has been little contact between the two bodies of though. This paper surveys cooperative game theory and explores the extent to which cooperative game theory may help us to understand (and perhaps extend) cooperative organizations. In particular, reciprocity motives are introduced into the cooperative game analysis, and this may provide a link between cooperative game theory and cooperative organizations.

In the sixty years since the publication of The Theory of Games and Economic Behavior, game theory has become an interdisciplinary field of research with applications in economics, political science, management, and philosophy, encompassing a wide range of models and methods. In particular, there is a broad dichotomy between noncooperative and cooperative game theory. Of the two, noncooperative game theory has been more successful and influential, if only in that the five men who have been honored with the Nobel memorial prize in economics for their work in game theory have all been honored primarily for work in noncooperative game theory. Nevertheless, the literature of cooperative game theory is extensive and important, and has some applications in economics and other fields. Cooperative game theory is applicable whenever the players in a game can form "coalitions," groups that choose a common strategy to improve the payoffs to the members of the group.

According to the International Cooperative Alliance, "A co-operative is an autonomous association of persons united voluntarily to meet their common economic, social, and cultural needs and aspirations …." There seems a clear parallel between these definitions. However, the definition of a cooperative organization continues "…through a jointly-owned and democratically-controlled enterprise." These elements do not occur in the concept of a cooperative game. Is there more to this parallel than the word "cooperative?" Has cooperative game theory anything to teach the Cooperative Movement?

This essay will begin with a historical introduction to cooperative game theory, and will then consider to what extent cooperative game theory should be considered a "positive" and to what extent a "normative" study. Some experimental evidence from the theory of *noncooperative* games will then be reviewed, and its implications for cooperative game theory are sketched. The experimental evidence suggests a richer view of human motivation than can be found in what we might call "classical" game theory, and a theory of cooperative games extended to take account of this evidence comes closer to the historic ideas of the cooperative movement. In this light, perhaps cooperative game theory can be useful in understanding and extending the cooperative movement.

The Emergence of Cooperative Game Theory

Much as economics springs from Adam Smith's *An Inquiry into the Nature and Causes of the Wealth of Nations,* game theory springs from a single influential book, *The Theory of Games and Economic Behavior,* by John von Neumann and Oskar Morgenstern. Von Neumann and Morgenstern adopted the mathematician's research strategy: to find a solution for the problem in its simplest form, and then to extend the solution step by step to more complex (and perhaps realistic) cases. Much later, Robert Aumann has suggested that a better name for game theory would be would be "interactive decision theory." The simplest cases of interactive decisions, in a clear sense, will be those with only two decision-makers, that is, two-person games. Von Neumann and Morgenstern simplified still further, considering two-person games in which anything gained by one must be lost by the other, that is, two-person zero sum games. Their solution to this simplest class of games – that each person will choose the strategy that maximizes his minimum payoff – is still recognized as the canonical solution of this simple case. But in a zero-sum game, there is no potential for cooperation in any sense, as there is (by assumption) no possibility of mutual gain. Thus, this was the beginning of *noncooperative game theory.* Later, John Nash would revise the von Neumann and Morgenstern solution in a form that could be extended to nonconstant sum games (that is, interactions with win-win and lose-lose potentialities) and demonstrate that all two-person games would have such solutions. This, with Alfred Tucker's famous Prisoner's Dilemma example, set the stage for the growth of noncooperative game theory.

However, this was not the direction that von Neumann and Morgenstern took. They observed that in any game with more than two players, or with win-win or lose-lose potentialities, players could benefit by forming groups for mutual benefit, that is, coalitions. They first studied games without win-win or lose-lose potentialities, that is, constant-sum games with 3 or more players. In these games the sum of the winnings of all players is constant, but it may be that one group of players can increase their earnings by "ganging up" on the others. Suppose that we can write down the total winnings that any subgroup can get, to divide among its members. The list of the total winnings for each group is called the "characteristic function" or "coalition function" for the game. A solution for a game in coalition function form should tell us 1) which coalitions will form, if any, and 2) how each coalition will divide its winnings among the members. In order to answer that, we must make some judgment as to what payoffs "rational" game-players will demand or expect or offer. The judgment that von Neumann and Morgenstern relied on was that a "rational" game player would not reject a take-it-or-leave it offer that is better than he could get individually in the absence of any cooperative agreement and (drawing on their solution for two-person zero sum games) that what he could get individually is the minimum payoff. Thus, von Neumann and Morgenstern suggested that a solution could be any efficient arrangement in which each individual got at least the minimum he could assure himself of by individual action. Understandably, in many cases there could be a very large number of solutions, and this was seen as a major drawback of the von Neumann and Morgenstern analysis.

In any case, they extended their analysis to nonconstant sum games by treating these games as having a "dummy" player. The "dummy" player does not play but has gains or losses that make up the losses or gains of all the "real" players. Thus, the nonconstant sum game is equivalent to a constant-sum game with one more player. In general, it would be to the advantage of the "real" players to gang up and extract as much as possible from the dummy, who, after all, has neither defenses nor pains of deprivation. Still, in general, there could be a very large number of solutions.

The ongoing research of cooperative game theorists in the following years mostly focused on narrowing the field – on refining or modifying the von Neumann and Morgenstern solution to reduce the number of possible solutions. Ideally the mathematician would like to have exactly one solution. Lloyd Shapley proposed a solution, the Shapley value, that had that desirable property. We shall have a bit more to say about the Shapley value later. Another proposed solution came from John Nash, who proposed a solution based on bargaining. Nash supposes that the bargainer will balance the risk of failing to come to an agreement against the benefit of demanding a higher share of the payoff. Like Shapley's solution, it had some attractive properties, and indeed has continued to be used in studies of bargaining; but it is difficult to extend Nash's bargaining theory to more than two parties in any common-sense way[1]. What neither Shapley nor Nash considered was this possibility: a subgroup might be able to organize themselves to gain total payoffs more than they are offered, and if so, would refuse the offer even if they could not do better individually. If we rule out all such offers of individual payments, we have the core of the game, an idea due to Gillies, and extended by Shapley, Martin Shubik, and others. This concept of solution seems to have been more widely applied in economics than the other cooperative solution concepts.

There are some other cooperative game solution concepts as well; we shall not go into them here. The number of different solution concepts is probably one reason why cooperative game theory has had less influence than noncooperative game theory. What they do all have in common is the assumption that people can make commitments to

[1] Harsanyi (1963) attempted to do so, but his contribution has not been much developed or applied in the subsequent literature, and in any case is presented as a modification of Shapley's value concept (p. 203 et. seq.)

carry out particular strategies in particular circumstances, in such a way that 1) the commitments are credible to the other players in the game, and 2) the commitments are carried out, even if the committed strategies are less advantageous than other strategies would be in the circumstances envisioned. The committed strategies may be cooperative in the sense that they make all parties better off than they would be without any agreement, while at the same time requiring that the parties refrain from opportunistic behavior. The committed strategies may also be threat strategies, designed to increase the player's bargaining power, but such that the threat is costly actually to carry out. Most noncooperative game theory assumes that people cannot make such commitments unless there is an external threat to carry them out, such as legal enforcement of a contract. However, contracts cannot solve all such problems. Much economic theory in the latter part of the twentieth century has made the same assumptions as noncooperative game theory, and holds that acting in accordance with commitments when other actions would have higher payoffs in the circumstances is *irrational*. On this view, if John Doe makes a commitment, and carries out the commitment despite the temptation to act opportunistically, and is better off as a result, John has acted irrationally. That is an interesting concept of rationality! On the whole, the difference between cooperative and noncooperative game theory seems to come down to a disagreement about *what is rational*; but the cooperative concept of rationality poses more difficult problems for the theorist. All concepts of cooperative solution hold that a rational person will always accept an offer that will leave him better off than he would be if he rejected it. They differ in their interpretation of when and whether the person will be better off to accept than to reject.

Some Ideas of Cooperative Game Theory

Even more than noncooperative game theory, cooperative game theory is a mathematical field, and relies especially on mathematical set theory for many of its basic ideas. The usual assumption is that *any* group of agents in the "game" can form a coalition, and a coalition among agents A, B, and C would be denoted as {A, B, C}. The brackets {} are conventional in set theory to indicate the "elements" of a "set," or in alternative ordinary language, the individuals making up a grouping. Now suppose A, B, and C coalesce,[2] that is, form a coalition. The expectation is that by working together and choosing a joint strategy they will be able to improve their results overall. It may be that one member, let us say C, bears a special cost for this, or another agent, such as A, gets most of the benefit. An example might be the modification of a river-course, so that those downstream benefit (with water supplies for irrigation) but those upstream lose (as some of their land is flooded). Then A is downstream and C is upstream. To enlist C in the coalition it may be necessary for A to pay C some compensation. It is common to assume "transferable utility[3]," which

---

[2] There is no commonly accepted verb form for "to form a coalition," but "to coalesce" seems apposite.

[3] This phrase recognizes that the ultimate benefits of economic activity are subjective, that is, in the economist's

means that a simple transfer of some of A's winnings to C can fully compensate C. In that case all that matters is the total payoff to the group {A, B, C}. Therefore, it is common in cooperative game theory to ignore all the details and to focus on the total values the various coalitions can obtain.

Most studies in cooperative game theory will then begin by an enumeration of all possible coalitions. Suppose there are N "players in the game." An individual agent can be indicated by A_i with i=1, … N. Common sense would see that any group of agents with more than one and less than N could form a coalition (with more or less difficulty). That's right, but it is not complete, and in cooperative game theory mathematical completeness is important. Therefore, in addition to those groupings, we also enumerate all singleton coalitions, { A_i }, that is, "coalitions" with just one member, and also the grand coalition of all N agents in the game and the null coalition, {∅}, a "coalition" with no members. (By convention in set theory ∅ means a set with no members). This mathematical completeness allows us to make mathematical statements about the game relatively concisely, and thus with less difficulty. Now, as suggested above, focus on the total value that each coalition can attain, and assign to each coalition a number expressing that value. It is commonly called the value of the coalition[4]. (Of course, the value of {∅} is zero). This assignment is called a "characteristic function" in mathematical set theory and is sometimes called the "coalition function" in cooperative game theory.

As von Neumann and Morgenstern noted, the problem is to find the numbers that correctly express "the total value that the coalition can realize." Suppose, for example, we have a large game consisting of two merchants and a large group of customers. If the merchants compete on price, the customers can realize benefits with a relatively good value of (let us say) ψ for each customer. Call the profits of the two merchants in this case π_1 and θ_1. In this example all are acting individually, that is, as singleton coalitions. Now suppose the two merchants form a coalition, that is, a cartel, and raise the prices to the consumers. The money value of consumer benefits are reduced to ν, less than ψ, while the profits of the merchants are increased to π_2 and θ_2. Now suppose the customers respond by forming a consumers' cooperative, that is, a coalition of all (or a large part) of the customers to supply their own needs rather than relying on the merchants. This raises the money value of their benefits to ζ, at least as large as ψ and

language, "utility," but assumes that utility is nearly enough proportional to money that money side payments will be sufficient to assure that everybody in a coalition benefits on net from the coalition's activity.

4 We are so accustomed to speaking of the value of an object that it may seem strange to speak of "the value" of a group of people. However, note that cooperators have long held that value is created (to use a common business phrase) by people working together, so speaking of the value of (created by) the group seems a common point between the cooperative movement and cooperative game theory.

perhaps greater, and reduces the profits of the merchants to zero[5]. Now, then, which of these values should we assign to these various coalitions? Following von Neumann and Morgenstern, we would assign the minimum – zero for a coalition of merchants and ν for customers.

An additional simplifying assumption that is usually made is that the game is "superadditive," that is, a coalition formed by the merger of two or more coalitions will realize a value at least as great as the sum of the values of the coalitions merged. We see this to some extent in the example: the merchants increase their profits by forming a cartel, and the customers increase their standard of living by forming a cooperative. Superadditivity would mean that a grand coalition of all merchants and customers together would have a value no less than $\pi_2 + \theta_2 + \zeta$ per customer. Since that is greater than the values of the coalitions separately, $0+0+\nu$ per customer, assigned as von Neumann and Morgenstern do, it seems that there is a large surplus that could be realized by a socialist ministry of distribution.

It should be noted that superadditivity is more than an assumption. As Aumann and Dreze (1974 p. 233) note, there are arguments for superadditivity that are quite persuasive, but, as they also note, superadditivity is quite problematic in many economic applications (including the one in the previous paragraph).

Von Neumann and Morgenstern were writing before Nash's work began the formation of noncooperative game theory. An alternative would be to assign the values of coalitions according to the Nash equilibrium in play among the coalitions[6]. Remarkably, this idea does not seem to have been developed until the 1990's (Zhao 1992). In general, market equilibria can be identified as instances of Nash Equilibrium. In that case, the values of the coalitions in the first instance – in which all agents operate as singleton coalitions – would be π_1 and θ_1 for the merchants and ψ for each of the customers. Notice, however, how the formation of a cartel by the merchants changes the value of a singleton consumer, and again how the formation of a cooperative changes the value of the merchants' cartel. These changes in the value of a coalition, as a result of the formation of another

[5] This example is suggested by *Paddy the Cope*, (Gallagher, 1942) which is still a very good read and a valuable biography of a life in the cooperative movement.

[6] It could be argued that there is an inconsistency in this, since Nash equilibrium assumes that commitments *cannot* be made as cooperative game theory assumes they *can*. One interpretation that reconciles them would be that the commitments can be made, but only within a coalition, while commitments to others outside the coalition, and to other coalitions, will not be honored if opportunistic behavior is more profitable. This is not unreasonable if we regard commitments as explicit promises and is consistent with a social norm of promise-keeping. In any case, since market equilibria are noncooperative equilibria in which coalitions interact noncooperatively, these assumptions are implied by ideas such as cartels, firms, employment relations, and cooperatives.

coalition, are called *externalities*[7] in recent cooperative game theory (Carraro). When the coalition function is taken as the only important information about the game, as so much cooperative game theory has done, this amounts to the unstated assumption that externalities are unimportant. This, too, seems quite problematic for economic applications. As early as 1963 Thrall and Lucas proposed a more complex way of assigning values to coalitions, the partition function, that allows for externalities in this broad sense. However, as Aumann and Dreze (1974 p. 233 note) remark, it creates technical (mathematical) difficulties. Perhaps for this reason it was not widely used in cooperative game theory until the 1990's.

Now return to the example of a coalition of {A, B, C} and suppose we have the values for each of the singleton coalitions, coalitions with just two members, and the value of the grand coalition v{A, B, C}. Taking the von Neumann and Morgenstern approach, the grand coalition will form (since the game is superadditive this will be efficient and therefore rational) and the only question is how to distribute its value among the members. Clearly no-one will settle for anything less than what he can get as a singleton, so the payoff to A, ψ_A, must be at least v{A}, and similarly for B and C. For efficiency, the sum of the payoffs ψ_A, must be the value of the grand coalition, v{A, B, C}. Nothing goes to waste. Von Neumann and Morgenstern then define a "dominance" relation to choose among payoffs schedule that may be stable. In general there can be many such payoff schedules. As we have seen, proposals by Nash and Shapley were designed to find a single solution among these many. Here is another way of narrowing the field. Suppose that A and B can secede together and realize a value of v{A, B}. They may insist that $\psi_A+\psi_B$ be at least v{A, B} to give them the incentive to remain in the grand coalition. Other subgroups who might secede can make similar demands. But suppose that v{A, B, C} < v{A,B}+v{C}. Then there is no way to assign all three agents payoffs large enough to persuade them to stay in the grand coalition, and we say that the coalition {A, B} *dominates* any payments that can be made by {A, B, C}. (In that case the game is not superadditive, but in more complex examples this problem can arise even in superadditive games.) If we narrow down the von Neumann and Morgenstern solutions to allow only those that are nondominated, we have the *core* of the game. But there may be no payoffs for any coalition sufficient to give the members reason to continue with the coalition, that is, the core may correspond to the empty set, ∅. On the other hand, as with the von Neumann and Morgenstern solutions, there may be many payoff schedules in the core. These are considered disadvantages of the core as a solution concept.

The Shapley value can always be computed and is unique, and has, as well, other attractive mathematical properties. The Shapley value assigns

[7] The term "externalities" is used more narrowly in economic theory. When a cartel is formed and thus imposes costs on customers, this would have been called a "pecuniary externality" in e.g Scitovsky; but modern economic theory does not regard "pecuniary externalities" as externalities.

payoffs to the members of the grand coalition and tells us nothing else about the joint action of coalitions. To obtain the Shapley value, first suppose that the agents in the game are approached in the order A, B, C. Each is offered his *marginal contribution*: that is, A is offered v{A}-v(∅)=v(A), B is offered v{A, B}-v{A}, and C is offered v{A, B, C}-v{A,B}. But this is somewhat arbitrary: the marginal contributions might have been different if they had been approached in a different order. Now suppose an arbitrator were to approach A, B, and C with the following proposition: "You can expect to receive your marginal contribution, but your marginal contribution depends on the circumstances, specifically the order in which you are recruited. Will you sign a contract that gives you the average of your marginal contribution over all possible circumstances, that is, all possible orders?" Being rational beings, but perhaps a little risk averse, we might expect them to accept. Anyway, this average of the marginal contributions is the Shapley value. It has the advantage that it is straightforward to compute (indeed a little easier than this example suggests in simple cases), and that it adds up to the total value, as well as the other advantages that have been mentioned.

These are the most widely used solution concepts. Let us digress by asking whether they represent "positive or normative" thinking.

Positive or Normative?

Economists are accustomed to drawing a distinction between "positive" and "normative" economics, a distinction due to Milton Friedman (1953). This distinction reflects a particular philosophy of science, logical empiricism, which is not universally accepted. Moreover, from a mathematical point of view, it is irrelevant. The same mathematical constructs may be used either for "normative" or "positive" analyses, as, for example, the Kuhn-Tucker Theorem in nonlinear programming can be used both in a "normative" model of maximization of a social welfare function and in a "positive" model of the substitution between inputs by a profit-maximizing firm. The validity of the mathematics does not depend on the purpose for which it is used. Despite these points, and with the caveats they imply, let us consider noncooperative and cooperative game theory in the light of the positive/normative dichotomy.

With respect to noncooperative game theory, the case seems pretty clear. Noncooperative game theory appears to be unambiguously "positive." While the same mathematics may be used either for positive or normative theorizing, it may not lend itself equally well to both uses. Generally, noncooperative solutions to games do not have any of the properties usually associated with normative theories in economics, and in particular, need not be efficient. In social dilemmas – generalizations of the well-known Prisoner's Dilemma example – and some other noncooperative games, inefficiency is one of the predictions of the theory. In a social dilemma two (or more) persons have to choose between two strategies, one relatively aggressive and one accommodative. While all are better off if all choose the accommodative strategy, each can be better off by opportunistically choosing the aggressive strategy. The empirical prediction of these noncooperative game models is that the aggressive strategy will always be chosen. This "positive economic" prediction is explicitly contrasted with the normative standard of efficient strategies.

The case is less clear with respect to cooperative game theory. In applications of social dilemmas, the accommodative strategies that maximize the total payoff are often described as "the cooperative solution" and represented as a normative goal that the noncooperative solution fails to realize. Indeed the accommodative strategies often called the "cooperate" strategies as opposed to the aggressive strategies. The aggressive strategies are called "defect" strategies. Implicitly, playing the aggressive strategies is seen as defection from the grand coalition of the players in the social dilemma. If we understand the social dilemma as a game with "transferable utilities," then the "accommodative" strategies do correspond to the cooperative solution in the broadest sense, since only the "accommodative" strategies provide an efficient outcome. Moreover, although studies of social dilemmas never explicitly analyze the "cooperative solution," and never consider side payments, the payoffs from both players playing accommodative strategies do realize the Shapley values for the game.

But the case is not quite that simple. Social dilemmas are a special class of games, and one of the assumptions that makes them special is an assumption that the players are symmetrical. Accordingly, applications of social dilemmas generally assume that the payoffs from cooperation are equally distributed. However, cooperative-game analysis does not require this. If one of the players could credibly threaten to "hold out" unless he were to obtain three-quarters of the gain from playing the "cooperate" strategies, that might well be a cooperative solution. But it will hardly be defended as a normative ideal.

If we look in more detail at the various cooperative-game solution concepts, it becomes more clear that none are really normative. The idea that links them all is that a rational, self-seeking person would not reject a take-it-or-leave-it offer for a deal that would give him a better payoff than the noncooperative solution. Since this usually leaves a large number of ways to divide the gains, the different solution concepts differ on the way that this is done. The von Neumann and Morgenstern solution set is agnostic, while the core assumes a kind of rugged collectivism, assuming that people cannot be denied what they can organize themselves to seize. This is not a normative perspective! Nash's bargaining theory supposes that people will determine their bargaining offers by balancing the risk of a breakdown in bargaining against the gain from having the demand accepted. This, again, would hardly be defended as a normative ideal.

Of the main cooperative solution concepts, the Shapley Value would be most credible as a normative concept, since it determines the payoff share of an individual by the overall contribution the individual makes to the coalition. It makes symmetrical payments to symmetrical players. Moreover, in the Prisoner's Dilemma and similar applications, where the "cooperative solution" is used loosely as a contrast to the noncooperative equilibrium, it is the

Shapley value that corresponds most closely to the "cooperative" strategies and payoffs. However, the Shapley Value can also be interpreted as an expression of bargaining power, and, indeed, has been used as an index of power. Thus it could be interpreted as a normative concept only if the "contribution" is construed in some normatively meaningful way, and it needs not be. The "contribution" may just be another way of expressing the individual's "power" over others. As we will see in the next section, the Shapley value can treat unsymmetrical players quite unequally.

Thus, none of the widely studied cooperative solution concepts is normative; all are positive and thus should be tested against the facts. But we have good reason to believe that none of these solutions corresponds closely to the actual world. All are, as we have said, based on the assumption that a rational human individual will never turn down an offer that improves on what he can get in the absence of a cooperative arrangement. Experimental evidence tells us that real human motivations differ from those assumed in economic theory and both cooperative and noncooperative game theory. With specific reference to cooperative game theory, there is evidence that people will sometimes refuse offers that would improve on the noncooperative solution of the game.

Reciprocity

The experimental studies in game theory have been primarily based on models in noncooperative game theory. One game that has been extensively studied is the "ultimatum game." (Henrich, e. g.) The Ultimatum Game is a two-person game along the following lines: the two agents may be able to share a fixed amount, such as $100. The first agent, the proposer, suggests a payment to go to the second agent, the responder. If the responder accepts the payment, he receives it, and the balance is paid to the proposer. However, if the responder rejects the payment, neither agent gets anything. The noncooperative equilibrium is one in which the proposer makes the smallest possible positive offer and the responder accepts it. However, experimental evidence disagrees with this prediction. If the proposer makes a very small offer, the responder is sometimes observed to reject the proposal despite sacrificing the small positive payment. Moreover, offers are often more than the minimum needed to avoid a rejection, and 50-50 offers are fairly common[8]. This appears to be an instance of reciprocity motives.

Traditional game theory proceeds from strong assumptions about human rationality to strong conclusions about the nature of equilibrium. One can ask whether either the assumptions or the conclusions are empirically valid, and indeed there is a long history of experimental studies in game theory that explore this point. Social psychologists and others quite early provided evidence that people facing a Prisoner's Dilemma-like game do not always act as neoclassical maximizers. (Lave 1965, Rapoport and Chammah, 1965, Morehouse 1967, Kreps et. al 1982).

---

[8] E. g. Guth , et. al., 1982, Henrich, et. al. 2005, and note also Roth (1995), Stanley and Tran (1998), Roth et. al. (1991), Andreoni and Blanchard (2006), Oosterbeck et. al. (2004)

Evidence of nonrational or non-equilibrium behavior in other games is equally plentiful.[9]

Some of the early experiments on the Prisoner's Dilemma were interpreted as evidence that altruism is an element in human behavior. Unfortunately, altruism is not always well-defined. Altruism was inferred, however, from a tendency to choose the cooperative strategy even when it is not a best-response strategy, for example, in Prisoner's Dilemma games. Some more recent studies have often focused instead on fairness or reciprocity. Berg, Dickhaut, and McCabe (1995) prefer the term "reciprocity," and say their "... results suggest that both positive and negative forms of reciprocity exist and must be taken into account ... [and] provide strong support for current research efforts to ... integrate reciprocity into standard game theory..." p. 139. In general, reciprocity motives take the form of negative and positive reciprocity, with positive reciprocity meaning that the agent intends to sacrifice her self-regarding interests to return a favor for a favor and negative reciprocity meaning that the agent intends to sacrifice her self-regarding interests to return a punishment for a wrong. Reciprocity motives cannot generally be characterized except by reference to norms or framing, consistently with what Gintis calls the strong reciprocity hypothesis. We would say that an act is framed as a favor when the other agent has contributed more than the norm and that the act is framed as a wrong when the other agent has contributed less than the norm. The ultimatum game provides a good example of the role that negative reciprocity can play. If the proposer makes a very small offer, the responder may perceive this as an aggressive act and sacrifice even the small payment he is offered in order to retaliate by refusing the offer, leaving the proposer with nothing. This is negative reciprocity. On the other hand, if the proposer's offer is generous, the responder may perceive this as a generous act. The responder sacrifices nothing by accepting and thus rewarding the generous offer. In other game experiments, however, a player might have to sacrifice something in order to reward the other's generous action. (For brevity examples will not be given here). This would be positive reciprocity.

Once again, let us continue with the example of the ultimatum game. As we have noted, there are several concepts of "the cooperative solution;" what they have in common is the idea that a person

[9] Because of the number of references -- without any representation of completeness -- it seems best to reserve them to a footnote. See Lieberman, B. (1960), Wolfe, G. and M. Shubik (1975), Guth, W and R. Schmittberger and B. Schwartze (1982), Guth, Werner and Peter Ockenfels and Markus Wendel (1993), Selten, Reinhard and Rolf Stoecker (1986), O'Neill, B. (1987), Rapoport, A. and R Boebel (1992), Roth, Alvin E. and Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir (1991), Cho, In-Koo and D.M. Kreps (1987), Diekmann, Andreas (1993), Berg, Joyce and John Dickhaut and Kevin McCabe (1995), Schotter, Andrew and Kieth Weigelt and Charles Wilson (1994), Roth, Alvin and Ido Erev (1995), Friedman, Daniel (1996), and Stahl, Dale O. and Paul W. Wilson (1995).

will never refuse a take-it-or-leave-it offer that makes him better off than the noncooperative alternative. The ultimatum game is especially suited to this question, since a take-it-or-leave-it offer is just what the responder faces. 1) Suppose, then, we follow von Neumann and Morgenstern in constructing the noncooperative alternative as the minimum payment. In that case *any* offer the proposer may make is a cooperative solution if it is accepted.[10] 2) Suppose instead that we construct the noncooperative alternative as modern noncooperative game theory does, viz. as the Nash equilibrium. Then the noncooperative payments are unique, and correspond to a minimum payment to the responder and the rest to the proposer. Since the proposer will not agree to a coalition that yields less than his noncooperative payoff, this is also the only cooperative solution. 3) Suppose we adopt the Shapley value and the minimal payment as the noncooperative alternative. Then the Shapley value for each player is just half of the total amount. By this criterion, the only cooperative solution is an equal split. 4) Suppose we adopt the Shapley value and the Nash equilibrium as the noncooperative alternative. Then the Shapley values are again the Nash equilibrium payoffs.

The point is that the experimental evidence clearly contradicts the last three of these interpretations of the cooperative outcome. For the first – the von Neumann and Morgenstern solution with their interpretation of the noncooperative outcome – the evidence is also negative in that 1) refusals do occur, and 2) extreme offers – close either to the Nash equilibrium or 50-50 – are less common than offers in between these limits. The Shapley value predictions on the two different approaches are clearly rejected. It seems that something is missing from all those hypotheses. That something may be reciprocity.

Continuing with the ultimatum game, suppose that the responder has internalized a social norm that a participant in such a game should receive at least x% of the payoff. An offer of less than x% will then be perceived as aggressive and ungenerous. (This perception is part of what we mean by the term “a social norm.”) Suppose, for example, that the norm is that each person should receive at least 25% of the amount to be distributed. Then the responder will reject offers substantially less than 25%. Conversely, the proposer, anticipating the responder’s acceptance (from which the proposer stands to benefit) may want to reward that friendly act (in advance) by offering something more than the norm of 25%, but the offer still will probably not approach 50%. These predictions would be similar for any social norm well within the interval from zero to 50%, and agree with the experimental results on the ultimatum

[10] Here as before I interpret von Neumann and Morgenstern as assuming that a rational being would be able to commit himself before the beginning of play to a contingent strategy along the lines of: “If I am offered at least x% then accept; otherwise refuse.” This transforms the ultimatum game into a version of Nash’s Demand Game (and indeed I conjecture that it transforms any two-person nonconstant sum game into Nash’s demand game). All possible payment schedules in the Nash Demand Game are Nash equilibria.

game. (An appendix, available from the author, gives more detail and a spreadsheet example.)

The ultimatum game is of interest primarily for its evidence of the importance of norms and reciprocity. Since it is a constant-sum game, it offers no possibility of a mutual benefit from working together. For cooperatives, and particularly worker cooperatives, the effort dilemma will be of more direct interest. This is addressed in McCain 2007.

What this suggests is that in order to understand cooperative *organizations* – and, as Gintis argues, a wide range of other social formations – we need to extend game theory with a broader conception of human motivation than has been part of game theory to date. Reciprocity motives and social norms would be a part of that conception. If we return to the general notion of a cooperative solution to the game as a "common strategy to improve the payoffs to the members of the group," social norms that support the common strategy would be among the dimensions of a "cooperative solution" of the game. To capture the idea, let us define the "extended cooperative solution" of a game as comprising 1) a common strategy that is efficient from the point of view of the members of the group, 2) a set of social norms that, in the presence of strong reciprocity motives, would result in the choice of the common strategy, without further enforcement, and 3) a set of guidelines for the sharing of the benefits of that common strategy, that is stable in the presence of a mixture of self-interest and reciprocity motives. The social norms for strategy choice and equable distribution are commitments, and to the extent that people are rational in the sense of cooperative game theory they will be carried out.

For an example of an extended cooperative solution in this sense, see McCain (2007) This is a "game" of effort determination in a cooperative enterprise, in which productive efficiency requires an effort commitment by each individual that is greater than the commitment the person would choose on the basis of pure self-interest in a noncooperative solution of the "game." Suppose that there is a norm of effort commitment, and suppose that the norm is the effort commitment required for efficient production. If one member makes effort less than the norm then other members perceive this as an unfriendly act and retaliate (perhaps by shunning or "putting in Coventry," though McCain is not explicit on this). Reciprocity motives lead to one or the other of two "solutions:" in one, everyone obeys the norm, production is efficient, and there is no retaliation. The other solution recapitulates the noncooperative one: everyone chooses the effort commitment on the basis of self-interest, without regard to social norms, and, each being equally a slacker, there is again no retaliation.[11] The "cooperative" norms, in this case, are the ones required for efficient production,

[11] In the noncooperative persective, both solutions are equally valid, and one implication is that the social norms will be effective in bringing about a cooperative solution only if there is a wide expectation that they will be followed, at least approximately. But we would hardly consider as norms a set of demands that no-one expects ever to fulfill or be fulfilled.

which also defines the common strategy. In this model, equal sharing of work time and pay are assumed throughout. As a result, one employee's effort is a favor to the other employees, and reciprocity motives lead them to reciprocate it. If, however, the enterprise is organized on the profit principle, then the benefit of one worker's increased effort goes to the proprietor, not to the other employees, and reciprocity motives would not lead them to increase their own effort, and the extended cooperative solution breaks down.

Of course, this is not a new idea. We find it in John Stuart Mill: "[C]ooperation tends ... to increase the productiveness of labour, … by placing the labourers, as a mass, in a relation to their work which would make it their principle and their interest -- at present it is neither -- to do the utmost, instead of the least possible, in exchange for their remuneration.' The key point for present purposes is the link it establishes among reciprocity, cooperative organizations, and cooperative solutions to games.

More generally, we began by observing that cooperative games and cooperative organizations have a common beginning in the idea of a group of people who choose a common course of action for their mutual benefit. The cooperative principles go beyond this by requiring that the group constitute "…a jointly-owned and democratically-controlled enterprise." It is the presence of reciprocity motives that makes those standards necessary, in many social contexts, to achieve a cooperative solution to the game. Minority ownership, and the remission of most of the benefits of efficient common action to the proprietor, in themselves seem to offend against reciprocity. As we have seen they interfere with the working of reciprocity among the workers in production. Monopolization of decision-making by a small minority also seems in itself to offend against reciprocity. Minority rule and profit may be stable and efficient so long as most people believe in the social norms expressed by the hymn: "The rich man in his castle, The poor man at his gate, He made them, high or lowly, And ordered their estate." That will be an easy norm for the rich man to believe in, but as ordinary people learn that they have alternatives, other and more equalitarian norms are likely to arise, and have arisen in most of the world.

Summary and Conclusions

Cooperative game theory and cooperative organization share the idea that agents join together and work together or choose a joint strategy for mutual benefit, but cooperative organizations are in addition jointly owned and democratically controlled. In the framework of cooperative game theory, profit-seeking firms (for example) are coalitions that manifest cooperative solutions just as cooperative organizations are. Like noncooperative game theory, cooperative game theory is best thought of not as a normative theory but as a positive or empirical hypothesis, a rational action hypothesis based on a slightly different concept of rationality than is noncooperative game theory. Indeed neither of these rational action hypotheses corresponds well with experimental evidence. However, when we modify the self-interest hypothesis of neoclassical economics and game theory, allowing for reciprocity motives and social norms, the experimental evidence can largely be accounted for. The reciprocity motives eliminate the most extreme predictions of self-interest theory in the experimental studies of the ultimatum game, and support the idea that in enterprise, a cooperative (in the sense of the cooperative movement) organization may be required in order to fully realize a cooperative (in the sense of game theory) solution to the interactive decision problem all group enterprises create.

References

Andreoni, James and Emily Blanchard (2006), "Testing Subgame Perfection Apart from Fairness in Ultimatum Games ," Experimental Economics v. 9, no. 4 pp. 307-321.

Aumann, Robert J. (2003), "Presidential Address," Games and Economic Behavior v. 45, pp. 2-14.

Aumann, R. J. and Dreze, J. H. (1974), "Cooperative games with coalition structure," International Journal of Game Theory v. 3, pp. 217-237.

Berg, Joyce and John Dickhaut and Kevin McCabe (1995), "Trust, Reciprocity, and Social History," Games and Economic Behavior v. 10, no. 1 pp. 122-142.

Carraro, Carlo (2003), The Endogenous Formation of Economic Coalitions (Cheltenham, U. K.: Edward Elgar).

Diekmann, Andreas (1993), "Cooperation in an Asymmetric Volunteer's Dilemma Game Theory and Experimental Evidence," International Journal of Game theory v. 22, pp. 75-85.

Friedman, Daniel (1996), "Equilibrium in Evolutionary Games: Some Experimental Results," Economic Journal v. 106, (Jan) pp. 1-25.

Friedman, Milton (1953), Essays in Positive Economics (Chicago: University of Chicago Press).

Gallagher, Padraic (1942), Paddy the Cope: An Autobiography (The Devin-

Adair Company, 2005 edition published by Kessinger Publishing.).

Gillies, D.B. (1953), Some theorems on *n*-person games (Ph. D. thesis, Department of Mathematics, Princeton University).

Gintis, Herbert (2007), "A Framework for the Unification the Behavioral Sciences," Behavioral and Brain Sciences v. 30, no. 1.

Guth, W and R. Schmittberger and B. Schwartze (1982), "An Experimental Analysis of Ultimatum Bargaining," Journal of Economic Behavior and Organization v. 3, pp. 376-388.

Harsanyi, John (1963), "A Simplified Bargaining Model for the n-Person Cooperative Game," International Economic Review v. 4, no. 2 (May) pp. 194-220.

Henrich, Joseph and Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Smith Henrich, Kim Hill, Francisco Gil-White, Michael Gurven, Frank W. Marlowe, John Q. Patton and David Tracer (2005), "'Economic man' in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," Behavioral and Brain Sciences v. 28, no. 6 (Dec) pp. 795-815.

International Cooperative Alliance (1995), "Statement on the Co-Operative Identity," available at: http://www.ica.coop/coop/principles.html, as of June 20, 2006, last modified on 2/9/2006.

Kreps, D. M. and R. Wilson (1982), "Sequential Equilibrium," Econometrica v. 50, no. 4 (Jul) pp. 863-894.

Lave, L. B. (1965), "Factors Affecting Cooperation in the Prisoner's Dilemma," Behavioral Science v. 10, pp. 26-38.

Lieberman, B. (1960), "Human Behavior in a Strictly Determined 3x3 Matrix Game," Behavioral Science v. 5, pp. 317-322.

McCain, Roger (2007), "Cooperation and Effort, Reciprocity and Mutual Supervision in Worker Cooperatives," Advances in the Economic Analysis of Participatory and Labor-Managed Firms edited by Sonja Novkovic (Greenwich, Conn: JAI Press).

Mill, John Stuart (1987), Principles of Political Economy (A. M. Kelley; Reprint of 1909 edition).

Morehouse, L. G. (1967), "One-Play, Two-Play, Five-Play and Ten-Play Runs of Prisoner's Dilemma," Journal of Conflict Resolution v. 11, pp. 354-362.

Nash, John (1950), "The Bargaining Problem," Econometrica v. 18, pp. 155-162.

Nash, John (1951), "Non-Cooperative Games," Annals of Mathematics v. 2, (Sep) pp. 286-295.

O'Neill, B. (1987), "Nonparametric Test of the Minimax Theory of Two-Person Zerosum Games," Proceedings, National Academy of Sciences v. 84, pp. 2106-2109.

Oosterbeek, Hessel and Randolph Sloof and Gijs van de Kuilen (2004), "Cultural

Differences in Ultimatum Game Experiments: Evidence from a Meta-analysis ," Experimental Economics v. 7, no. 2 (June) pp. 171-188.

Poundstone, William (1992), Prisoner's Dilemma (New York: Doubleday).

Rapoport, A. and R Boebel (1992), "Mixed Strategies in Strictly Competitive Games: A further Test of the Minimax Hypothesis," Games and Economic Behavior v. 4, pp. 261-283.

Rapoport, Anatole and Albert M. Chammah (1965), Prisoner's Dilemma (University of Michigan Press).

Roth, Alvin and Ido Erev (1995), "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," Games and Economic Behavior v. 8, pp. 164-212.

Roth, Alvin E. and Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir (1991), "Bargaining and Market Behavior in Jerusalem, Ljubliana, Pittsburgh, and Tokyo: An Experimental Study," American Economic Review v. 91, no. 5 (Dec.) pp. 1068-1095.

Roth, A. (1995), "Bargaining Experiments," Handbook of Experimental Economics. edited by Kagel, J., Roth, A. (Princeton: Princeton Univ. Press) pp. 253-348.

Schotter, Andrew and Kieth Weigelt and Charles Wilson (1994), "A Laboratory Investigation of Multiperson Rationality and Presentation Effects," Games and Economic Behavior v. 6, pp. 445-468.

Selten, Reinhard and Rolf Stoecker (1986), "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames," Journal of Economic Behavior and Organization v. 7, pp. 47-70.

Shapley, L. S. (1953), "A Value for n-person Games," Contributions to the Theory of Games II edited by H. W. Kuhn and A. W. Tucker (Princeton: Princeton University Press) pp. 305-317.

Shapley, L. S. and Shubik, Martin (1969), "On Market Games," Journal of Economic Theory v. 1, pp. 9-25.

Smith, Adam (1937), The Wealth of Nations (New York: The Modern Library).

Stahl, Dale O. and Paul W. Wilson (1995), "On Players' Models of Other Players: Theory and Experimental Evidence," Games and Economic Behavior v. 10, pp. 218-254.

Stanley, T. D and Ume Tran (1998), "Economics Students Need Not Be Greedy: Fairness and the Ultimatum Game," Journal of Socio-Economics v. 27, no. 6 pp. 657-663.

Thrall, R. M. and W. F. Lucas (1963), "n-person games in partition function form," Naval Research in Logistics Quarterly v. 10, pp. 281-298.

von Neumann, John and Oskar Morgenstern (1944), Theory of Games and Economic Behavior (Princeton: Princeton University Press).

Zhao, J. (1960), "The Hybrid Solutions of an N-Person Game," Games and Economic Behavior v. 6, no. 3 pp. 145-160.