Determinacy and Rational Choice

Russell Hardin
Departments of Political Science and Philosophy, University of Chicago
5828 South University Avenue, Chicago, IL 60637, USA

Suppose we have a fully determinate rational choice theory that will tell each of us what to do in a particular interactive choice context and suppose each of us knows the positions of all the others. From our theory I can calculate not only what I ought to do but also what each other agent in the interaction ought rationally to do. Given knowledge of what each other agent should do, I can check whether I could in fact do better by not following the theory. If that theory yields determinate solutions, in the sense that it tells each of us explicitly what to do, these solutions must be in equilibrium for rational players who have full knowledge of the payoff structure, who have this theory, and who assume that their co-players are rational and have this theory.

If a purported solution were not in equilibrium, then, by definition, at least one of the players would have an incentive to choose a different strategy in order to receive a better payoff. But a coherent theory cannot be used against itself to tell me to do other than what it commends. Hence, each of us must find that what the theory tells us to do is, once all others are also following the theory, an equilibrium outcome.

John Harsanyi has long argued for determinacy in game theoretic rational choice theory (Harsanyi 1956, 1977) and, not surprisingly, he has also generally argued for equilibrium outcomes as rational, as in his recent work with Reinhard Selten (Harsanyi and Selten 1988). Against this general position, I wish to argue that primitive notions of rationality do not entail determinacy when followed by all agents to an interaction. The imposition of determinacy as a part of the notion of rationality is an extra-rational move that is not compelling in many contexts. In particular, it is not compelling in contexts, such as iterated Prisoner’s Dilemma, in which our theory must give contingent instructions on how to play into the future.
Rationality

First, focus on the underlying intuition that stands firm through all the problems we may face: A rational person prefers more value to less or, absent value assignments, has transitive preferences. For present purposes, I will assume cardinal utility in keeping with the overwhelming practice in game theoretic discussions, so that we may dispense with reference to transitivity in a strictly ordinal theory of preference. In cardinal contexts, preference for x over y merely means the value of x is greater than the value of y. Many critics of this notion of rationality make the simplistically one-step, materialist mistake of supposing the principle implies one should prefer more to less of anything valued at all. But, obviously, I may not prefer more dinner to less tonight because I may not prefer the consequences of still more once I have consumed enough. Indeed, I am unlikely even to have appetite for more beyond some level of consumption. Cardinal preference for more is strictly preference for more value or utility and is therefore trivially transitive.

If my preferences are transitive in this cardinal sense, it follows that, if I am choosing from n possibilities, there must be one or more possibilities that I prefer to all others, or that I hold indifferent to all, so that in this case preferring more to less value is equivalent to maximizing. Our more general problem is how to choose under varied constraints when we have to compromise with the principle of maximization. An obvious a priori case in which there can be no compellingly unique compromise is that in which I am offered a choice from an infinite set of possibilities none of which has the greatest value (e.g., $1000, $2000, $3000, ...). But this, unfortunately, is not a practical problem.

Maximization is commonly treated as the basic meaning of rationality or, at least, as part of the meaning. In particular contexts we have no difficulty giving clear meaning to the notion of maximizing in individual choice over various alternatives. For example, given a choice between two outcomes, one of which I prefer to the other, I know what it means to maximize. If my choices or actions determine results only probabilistically, the notion of maximization is already in question. Most choice theorists seem to find the generalization to maximum expected utility an obvious extension of the primitive notion of rationality, although some people do not find that move obvious or even right. The central problem of rationality in economics is that we cannot so readily or agreeably generalize our notion of maximization to many other contexts. Most especially, we cannot generalize it to cover contexts in which choices of different agents interact in producing joint outcomes for the agents. The theory of games was invented to model
such choice problems.

The difficulty in generalizing our notion of maximization to cover all games is that each agent is to maximize his or her own utility function over all the outcomes while different agents may have quite different utility functions over those outcomes. It is not always impossible to maximize over two functions at once, but it is not possible to do so in general or to specify a general rule for doing so. A trivial case in which genuine multiple maximization is possible is when our maxima are all in the same outcome even though our functions are not otherwise identical, as in certain coordination games. Despite special cases such as these, however, it is still true that there is mathematically no solution to the general problem of multiple maximization. Hence, much of game theory, and of choice theory and economics more generally, is directed at finding compromises to multiple maximization that seem to be what one would deduce from the principle of individual maximization. As compelling as some such compromises might be in certain contexts, they are still not as compelling as maximization over a single function that has a greatest value at some point or points.

The first success of game theory was Neumann’s saddlepoint theorem for two-person constant sum games. In such a game, the two players have inversely related utility functions over the outcomes. Clearly, it is impossible to maximize both functions at a single outcome. But, because the game is constant sum, these two functions are not independent, so that there might be a plausible compromise to simultaneous maximization, maximization subject to a constraint. In the two-person constant sum game both players face the same constraint, defined by the inverse of each player’s own utility function over the outcomes. Neumann’s maximin strategy seems compelling to virtually every game theorist who has considered the problem. One may say it is a compelling compromise that seems consistent with what we would expect from maximizing players, given that maximization is mathematically impossible for both at once. But the description of that solution — maximin — is not a good prescription for a solution to the general case in which the functions over which we are simultaneously trying to maximize may be independent so that we do not face analogous constraints.

There are other contexts in which multiple maximization — that is, maximization over more than one function at once — seems to have compelling implications. In a market of many sellers and many buyers, whose utility functions over all possible outcomes of trade cannot be simultaneously maximized, we may nevertheless expect a very restricted range of possible outcomes to result. Again, although the argument may be a bit convoluted, this result follows from what seems similarly to be a compelling
compromise implicit in or consistent with the notion of individual maximization. All players in the market face the constraint that large numbers with similar interests typically imply the failure of collective action to collude on prices for sale or purchase.

Equilibrium

As it happens, all of the results above — those that follow from literal maximization and those that follow from some compromise more or less deducible from it — are equilibria in their particular games or interactions. For many contemporary economists that is not surprising, because they suppose equilibrium is inherently bound up in rationality. But the bonds go in only one direction. Equilibrium is defined in rational terms; rationality is not defined in terms of equilibrium. An equilibrium is an outcome from which no one would have incentive to deviate, by choosing a different strategy, even if it were possible to second-guess all the others in the interaction after knowing what choices they have made.

That the bonds do not go the other way can be seen from the facts that (1) there can be multiple equilibria with differing payoffs to various players, and (2) there can be equilibria that are disastrous for all concerned and that are Pareto inferior to other possible outcomes. In both cases, it may be true that some subgroup has an interest in switching collectively out of the strategies that put them in the equilibrium. This move is often prohibited in individualistically defined equilibrium theories. Yet, much of the social interaction we wish to explain is just such subgroup cooperation for the subgroup's interest. In their solution theory, Harsanyi and Selten resolve many of the cases under (2) by invoking their principle of “collective rationality,” according to which, “in the absence of special reasons to the contrary, rational players will choose an equilibrium point yielding all of them higher payoffs, rather than one yielding them lower payoffs” (Harsanyi and Selten 1988, p. 356).

Equilibrium is defined in terms of individual rationality. If we are in equilibrium, then it would not be rational for any of us individually to choose a different strategy. Frank Hahn has suggested an alternative notion of general economic equilibrium that makes the bonds go both ways. We may restate his notion to fit game theoretic choice problems: Agents in an interactive choice are in equilibrium if their choices generate messages that do not cause agents to change the theories they hold or the policies they pursue. Hahn says the central point of this conception of equilibrium is that one abandon
one's theory when it is “sufficiently and systematically falsified” (Hahn 1984, p. 59). At first reading, this may sound like an empiricist view. I wish to read it, rather, as a theoretical view.

Robert Aumann (1985, pp. 28-9) says that a game theoretic “solution concept should be judged more by what it does than by what it is; more by its success in establishing relationships and providing insights into the workings of the social processes to which it is applied than by consideration of a priori plausibility based on its definition alone.” Selten (1985, p. 81) replies that: “Descriptive theories need to be compared with reality, whereas normative theories cannot be tested empirically. The justification of normative theories of rationality must be sought in compelling intuitive arguments.” Similarly, Richard Jeffrey (1983, p. 166) notes that the Bayesian theory of preference is ours — the theorists’ — but not necessarily the agent’s. Such a theory is, in this respect, not an analog of, say, the neurological theory of vision, which is also not the agent’s theory. Rational choice theory is not descriptive, as the neurological theory is, but normative. This means, incidentally, that the normative theory of rational choice is very hard — indeed, it is beyond most people including theorists.

As I understand them, Selten’s and Jeffrey’s view on this issue is far more compelling than Aumann’s. In what follows, I will argue entirely from intuitive tests of arguments. In particular, in keeping with Hahn’s notion of equilibrium, I will test Harsanyi’s determinacy principle for game theoretic solution theory. There are, in any case, virtually no data on real-life, many-times but finitely iterated Prisoner’s Dilemmas.

Iterated Prisoner’s Dilemma

Few problems in game theory have generated as much discussion as the iterated Prisoner’s Dilemma, especially when it is iterated a fixed number of times. Very early it was proposed that, if it is rational to defect in a single-play Prisoner’s Dilemma, then it is also rational to defect in every play in a fixed iteration of plays of the game. The argument, from so-called backwards induction, is that it must be rational to defect on the last play because there are no longer any incentives from future plays to bring into consideration. But then that last play cannot provide incentives in prior plays and both players must know this to be true, so that it must be rational to defect on the next to last play. And so on back to the first play. (Formal demonstrations of this transparent result abound [e.g., Friedman 1986, pp. 95-6].) There have been many practical and theoretical
objections to that conclusion and there have been many and varied proposals for seemingly more sensible resolutions (Luce and Raiffa 1957, pp. 97-102; Radner 1980; Hardin 1982, pp. 138-54; Sorenson 1988, pp. 344-61).

Determinacy yields an argument for the rationality of continuous defection that is much simpler than the traditional backwards induction. From the symmetry of the players, a determinate theory must prescribe the same strategy to each player. We may suppose it permits mixed strategies, so long as they are well defined. If it prescribes cooperation on the final, nth \((n >> 1)\) play of the series, either player can do better by violating the theory. Similarly, suppose that, on the final play, it prescribes a probabilistic mix of strategies, \(pC + (1 - p)D\), where \(C\) is cooperation, \(D\) is defection, and \(1 \geq p > 0\). Then either player can do better by defecting on that play. If it prescribes defection on all plays after some number \(q\) \((n > q \geq 1)\) with cooperation or a mix of cooperation and defection on play \(q - 1\), then either player can do better by defecting on play \(q - 1\). Hence, the theory cannot prescribe either cooperation or a mix of cooperation and defection on the final play and it cannot prescribe either of these on any play that precedes a final series of one or more defections. It must therefore prescribe defection on every play. Hence, either we should defect on every play in a fixed iteration Prisoner’s Dilemma or we can have no acceptable determinate theory.

Against the first half of this conclusion, many have noted that the players in a Prisoner’s Dilemma that is iterated many times could do better than the mutual always-defect outcome. That is just the core meaning of rationality: to choose more rather than less. Already Luce and Raiffa (1957, p. 100) supposed that rational players would cooperate for the first 90 or so plays of a 100-play Prisoner’s Dilemma. Unfortunately, it is not clear what it would mean to maximize in this context. Nor may anything like the Luce and Raiffa solution be in equilibrium under a given equilibrium concept. Of course, the always-defect strategy also does not maximize. We could say that a particular strategy would maximize only against some particular strategy followed by the other player. If that other player’s strategy were determinate, then I could know how to maximize. If we constrain rational choice to yield only the always-defect strategy, then it follows that each can maximize with respect to that strategy by playing it.

To deal with such problems as the apparent irrationality of always-defecting in a finitely iterated Prisoner’s Dilemma, game theorists have proposed many equilibrium concepts to single out better results than the sometimes dismal Nash equilibrium. These concepts include proper, perfect, subgame perfect, sequential, and persistent equilibria. These are not equivalent and some may fairly be called arcane. All of the debate here is
over what counts as a good rational choice theory.

Given that we — rational choosers — have difficulty finding and agreeing on such a theory, we should hedge our own rational choices by considering the possibility that the theory we hold at the moment will eventually seem wrong. If my theory says I should defect continuously in iterated Prisoner’s Dilemma, I should hold open the possibility that you will convince me of the superiority of another theory that recommends otherwise. Hence, when we begin playing our many-times iterated Prisoner’s Dilemma, if you cooperate on the first or some early move, I should reconsider my theory as Hahn suggests. Of course, because of the symmetry of our reasoning under any correct theory, I should also consider what would be the effect of my own cooperative choice on the first play (Hardin 1982, p. 149; for related but more complex views, see Bicchieri 1988). If you are, say, twice as modest as the average academic game theorist, you will likely reconsider the wisdom of continual defection once I have cooperated and thereby wrecked your expectations. Here we do not invoke the trembling hand, which might mistakenly choose the wrong strategy. Rather, we invoke a trembling intellect that, rightly, has reason to doubt its own theory. In this case, by violating my own theory, I can expect to improve my own long-term payoff. But that just means my theory is self-contradictory. If a theory recommends its own violation, it is not in equilibrium under the Hahn notion of equilibrium. The theory not only trembles, it collapses.

Where does this leave us? Evidently, we are stuck with indeterminacy in game theoretic solution theory and in rational choice theories more generally. Our theory cannot tell players in an iterated Prisoner’s Dilemma fixed to run, say, 100 plays that they should cooperate for most of the plays and then start defecting at the 98th play even if the other is still cooperating on the 97th play. It can only say that efforts at cooperation may be individually rational until far along, perhaps until nearly the final play of the series. If the series is short — say, two or three plays — the theory might commend defection all the way if it commends defection in the single-play game. (And, of course, if the payoff for cooperating while the other defects is very large relative to the difference between the both-cooperate and the both-defect payoffs, it may not be rational to cooperate at all.)

In single-play Prisoner’s Dilemma, there is an additional consideration that makes defection seem compelling: Defection dominates cooperation, which is the only other pure strategy. (It therefore also dominates every mixed strategy.) In iterated Prisoner’s Dilemma, always-defect does not dominate all other pure strategies. It may be said to chain-dominate them, as in the backwards induction argument. For example, Tit-for-tat
all the way is dominated by Tit-for-tat with defection on the last play (call this T). The latter is dominated by Tit-for-tat with defection on the last two plays (call this T - 1). And so forth. But chain dominance is not a compelling principle for choice. It leaves open the possibility that, for example, T - k is inferior to T. In pairwise choices over various strategies, always-defect might lose to many other strategies. In any case, there are many strategies which it does not dominate. To force its selection by imposing determinacy or equilibrium conditions on final outcomes goes well beyond the primitive notion that it is rational to choose more rather than less value or to have transitive preferences.

The sure thing or dominance principle sounds misleadingly like a transitivity requirement over strategy choices (not outcomes). Strategy A dominates strategy B, we eliminate B. Now among the strategies in the newly reduced set, C is the strategy our theory selects. But this does not entail transitivity of dominance or even of strategy rankings. If B were still in the set, it might be chosen over C because it beats C for some strategy choices of other players or because it would elicit more beneficial strategy choices from others. Under our theory without the elimination of dominated strategies, we prefer C to A to B to C.

If A, B, and C were states of affairs or values, we would want our preferences over them to be transitive. Does it make sense to have intransitive preferences over strategies? It would be wrong to have intransitive preferences over strategies if the strategies determine outcomes directly, for example, if all other players’ strategy choices are already known. But this is because, under these conditions, intransitive rankings of strategies entail intransitive preferences over states of affairs.

More generally, however, it is odd even to speak of preferences over strategies. What we normally speak about is preferences over alternative states or over different values. Then we may rank strategies according to the payoffs we expect them to bring us. My choice of any strategy will bring me a return that is a function of what strategy you choose. To determine what I expect my return to be in an iterated interaction, I must estimate what you will likely do in response to my strategy choice. Then I may say in the ordinary way that I prefer the likely state of affairs that follows from my choosing one strategy to the likely state of affairs that follows from my choosing the always-defect strategy. In a many-times iterated Prisoner’s Dilemma with reasonably favorable ratios of the payoffs, one suspects that almost everyone would choose to play cooperatively to some extent to engender cooperation.

A surprising number of recent equilibrium concepts seem to have been motivated at least partly by the compelling intuitive sense that the always-defect strategy in finitely
iterated Prisoner’s Dilemma is not rational. To suppose it is rational is to suppose that very good game theorists would play always-defect for hundreds, thousands, or more plays if only the trial is to terminate after a set number of plays. That is an absurd supposition. This is only an intuitive claim for cooperation, but it is a firmer intuition than any intuition that rational solution theory must be determinate or must produce standard equilibrium outcomes in ongoing contexts. It appears to be an intuition shared by many, perhaps most, contemporary game theorists, who seem to be motivated by the evident failure of the standard equilibrium argument of backward induction.

Concluding Remarks

Perhaps the demand for determinacy comes from earlier simplifying assumptions of complete and perfect information. If we all have complete information about what all others are going to do, and they have complete information about us, the outcome of our interactive choices must be determinate, at least as from now. But that means it must be an equilibrium. If it were not, at least one of us would have incentive to act in some way other than that in which complete and perfect information says he or she will act. One might suppose that this is paradoxical and therefore conclude that the assumption of complete information on future choices is an incoherent notion. (This is one response to some ways in which Newcomb’s problem is stated [Hardin 1986].) Alternatively, one might conclude that the notion is not incoherent, that there are no actual choices to be made just because all actions are determinate. This move raises the usual conundrums of determinism in philosophy. Not the least troubling of its implications is that choice theory cannot be prescriptive because there are no agents to whom it could prescribe anything.

Whatever motivates that demand, however, requiring determinacy in rational choice theory is less compelling than the urge to prefer more to less value in our interactions. To require determinacy sets up the cute claim that it is rational to act irrationally in iterated Prisoner’s Dilemma, because it is beneficial to act cooperatively despite the rational dictum always to defect. As Harsanyi notes, for game theory we must go beyond the rationality postulates of individual decision theory (Harsanyi, 1977, p. 10), or we cannot resolve our problems. Iterated Prisoner’s Dilemma seems to provide a major litmus test for the general applicability of various extensions of the notion of rationality. Determinacy fails that test.
References


