NEAL C. BECKER AND ANN E. CUDD

INDEFINITELY REPEATED GAMES: A RESPONSE TO CARROLL

ABSTRACT. In a recent volume of this journal John Carroll argued that there exist only uncooperative equilibria in indefinitely repeated prisoner's dilemma games. We show that this claim depends on modeling such games as finitely but indefinitely repeated games, which reduce simply to finitely repeated games. We propose an alternative general model of probabilistically indefinitely repeated games, and discuss the appropriateness of each of these models of indefinitely repeated games.

Keywords: Repeated games, prisoner's dilemma.

In his article 'Indefinite Terminating Points and the Iterated Prisoner's Dilemma', John Carroll asserts that there are "only uncooperative equilibria in finitely, but indefinitely, iterated games," and that this "calls into question the significance of the existence of cooperative equilibria in infinitely iterated Prisoner's Dilemma games."[1] In this note we will show that the second claim is overstated because the iterated games he refers to are equivalent to finitely and definitely iterated games, of which it is well known that there are only uncooperative equilibria. We will show that there exists an alternative model of indefinitely iterated games which is not equivalent to finitely iterated or to standard infinitely iterated games, but which allows for cooperative equilibria in some instances. And we will argue that these games capture the features of reality which Carroll wanted to highlight.

The problem centers around Carroll's notion of an indefinite iteration, which he characterizes using what he calls a 'terminating **p**-function', defined as follows.

p is a terminating **p**-function if and only if

(1) $\sum_{t=1}^{\infty} p(t) = 1$, and

(2) there exists a natural number λ such that

Theory and Decision **28**: 189–195, 1990.

(a) $\mathbf{p}(\lambda) > 0$, and

(b) for all natural numbers $n > \lambda$, $\mathbf{p}(n) = 0$.[2]

He tells us that we can think of $\mathbf{p}(t)$ as the probability that the game will be iterated t times. Notice that according to this definition the game will not be played more than λ times, so that the sense in which the endgame is indefinite is that it might be played fewer than λ times. Since there is an iteration which will be the last if the game is iterated even that long, it seems to us that it is not at all surprising that a form of the backwards induction proof that shows that there are only noncooperative equilibria in finitely repeated Prisoner's Dilemma (hereafter, PD) games would apply here, as Carroll demonstrates.

We will argue that the indefinite iteration to which Carroll refers is just equivalent to a finitely repeated game with a definite final iteration. This would show what we claim, which is that Carroll's result is a trivial extension of an already well-known result. Suppose we have the following PD payoff matrix:

		player 2	
		C	D
player 1	C	x, x	z, y
	D	y, z	w, w

Fig. 1.

The payoffs are real numbers related as follows: $y > x > w > z$. Carroll defines a function $\mathbf{p}^*(t)$ to represent the probability that the game is iterated in period $t + 1$.[3] Let ${}_{\langle a,b\rangle}\mathbf{X}_{i,t}$ be the payoffs to player i in period t given that i plays supergame strategy a and the other player plays supergame strategy b. The iterated supergame payoff is then given by the sum of the payoffs in each game given the strategy pairs of the players multiplied by the probability that they will be played:

(1) $$\sum_{t=1}^{\infty} p^*(t-1)_{\langle a,b\rangle}X_{i,t}.$$

Note that this is just equivalent to:

$$(1') \qquad \sum_{t=1}^{\lambda} p^*(t-1)_{\langle a,b\rangle} X_{i,t},$$

since λ is the last period with a positive probability of being played. The '∞' in (1) is gratuitous, and serves only to cloud the issue of whether the game is to be repeated finitely but indefinitely (i.e. with a finite upper bound on the number of possible iterations) or infinitely but indefinitely (that is, with no finite upper bound on the number of possible iterations). Note also that the utility functions which Carroll is presupposing must be von Neumann–Morgenstern utility functions, since he assumes in writing the iterated game payoff functions this way that they have the expected utility property.[4] Finally note that Carroll seems to be supposing that the players have common knowledge of the iterated supergame, in particular, of λ, and of the players' rationality.[5]

The problem, then, is to find the equilibrium strategies a and b for each player, given the supergame payoffs in Equation (1′). By the expected utility property, the problem is equivalent to finding the equilibrium strategies a and b for the supergame payoff given by:

$$\sum_{t=1}^{\lambda} {}_{\langle a,b\rangle}\mathbf{X}^*_{i,t},$$

where $\mathbf{X}^*$ represents the payoffs of the following matrices for each period:

		player 2: C	player 2: D	
player 1	C	$\mathbf{p}^*(t-1)(x, x)$	$\mathbf{p}^*(t-1)(z, y)$	Period t
	D	$\mathbf{p}^*(t-1)(y, z)$	$\mathbf{p}^*(t-1)(w, w)$	

Fig. 2.

But this is just λ, i.e. finitely many, consecutive *PD* games. In the last iteration players have a dominant strategy to play D, as in the one shot

PD, and this begins the backwards induction. Thus Carroll's indefinitely iterated *PD* supergame is equivalent to the definitely finitely iterated *PD* supergame with discounting for time, and so it is no surprise to learn that it has only an uncooperative solution.

It is still an interesting question, however, whether situations of indefinitely repeated interaction should be modeled as finitely repeated discounted games or as the standard sort of indefinitely repeated games. We would like to conclude by giving a few reasons to think that the indefinitely repeated variety is better suited to the task of modeling a risky future. The 'standard indefinitely iterated game', as we shall call it, is the game represented by Figure 1 repeated with probability d each time, so that the probability that the game will be repeated once (i.e. played twice) is d, $(0 < d \leq 1)$, the probability that it will be repeated twice (i.e. iterated 3 times) is d^2, and in general, the probability that iteration t of the game, g_t, will be played is d^{t-1}. The difference between this and Carroll's finite but indefinitely iterated game is that the standard iterated game has a small but positive probability of continuing in any finite iteration. But since it is indeed certain that we are all mortal, and thus face only finitely many iterations of any decision situation, Carroll claims that the iterated game with a definite upper bound on the number of iterations is a better model of reality. He writes, "my definition of the iterated payoff is specifically designed to capture the finiteness of genuine iterated Prisoner's Dilemmas."[6]

There are two problems with Carroll's claim. First, as we have shown, the definiteness induced by the upper bound on the number of iterations in Carroll's game (i.e. clause (2) of his definition of the terminating p-function), overwhelms any indefiniteness in the model, and makes it equivalent to a finitely and definitely iterated game. That is, by introducing a definite endpoint Carroll has trivialized the indefinite nature of the iterations. Second, the upper bound on the iterations fixes the conditional probability of future iterations of the game beyond λ. In particular, the conditional probability of playing another game given that one is playing the λth iteration is 0. But is it realistic to suppose that one is usually, or perhaps ever, certain that the current interaction is the last one? We think that it is more realistic to imagine that in any interaction there is a positive probability that one will experience another interaction; that's why it is a bad idea (unless you are precommitting yourself to preclude weakness of will) to burn your bridges.

An analogy may make our point clearer. Suppose Ethel lives longer than anyone has ever lived before, say 175 years. We might imagine that a game theorist modeling *PD* iterations in life would set the probability of living 176 years at 0. But is it reasonable for Ethel to believe that it is certain that she will not live another year? Given that she has lived 175 years, it seems to us that she would be justified in placing a positive, if small, probability on living another year. The same point could be made, we believe, for any interaction situation – it seems to us that it is normally unreasonable to believe with certainty that any particular interaction is the last of its kind. Carroll claims that it is certain that we will not face a decision situation for millions of centuries, but in order to build this *certainty* in he must set an arbitrary upper bound on the number of iterations, and this is problematic for at least the last game. The standard indefinitely iterated game, on the other hand, would place a very small, vanishing to zero, probability on the chances of interactions continuing for "millions of centuries".

An alternative approach, which captures the idea that the probability of continuing iterations of the game decreases with time, but doesn't require the conditional probability of the game continuing at any stage to be zero, is to let the probabilities decrease monotonically with time reaching zero only in the limit, if at all. Let $p'(t)$ be the probability that the game will be played t times conditional upon its having been played $t-1$ times, such that p' has three properties:

(*a*) $\lim_{t \to \infty} p'(t) = \alpha, \qquad 1 \geqslant \alpha \geqslant 0;$

(*b*) $p'(t)$ is monotonically decreasing;

(*c*) $1 \geqslant p'(t) > 0.$

Notice that if we were to allow $p'(t) = 0$ for some t, we would have Carroll's finitely iterated *PD*. If $p'(t) = 1$ for all t we have the infinitely repeated *PD*, and if $p'(t)$ is a constant between 0 and 1 we have the standard indefinitely repeated *PD*. Our $p'(t)$ function also allows the conditional probability to vary with time in order to capture the situation of Ethel.

For what we have called the standard indefinitely iterated *PD* (i.e. the infinitely repeated game with constant discounting) it is well known that for α small enough (relative to w, x, y, z) there exists a finite iteration t^* such that the probability of continuing is very close to α and it no longer pays the players to cooperate, and this is enough to begin the backwards induction to show that they will never cooperate for such an α. Thus for some such games Carroll's result would hold despite the potentially infinite iteration. More interestingly, however, Fudenberg and Maskin (1986) showed that if the players are sufficiently patient (or equivalently the probability of continuing is sufficiently great) in the indefinitely iterated *PD* there are also equilibria at which players cooperate.[7] (In our $p'(t)$ function the sufficient patience requirement is that α must be large enough.)

Of course, there may be artificial situations which are best modeled by the indefinitely finitely iterated *PD*, such as an Axelrod-type tournament in which an upper bound on the number of iterations has been set. But these situations must be manufactured to have the definite upper bound; they are not naturally occurring. And as we have seen, such games are equivalent to finitely repeated games with discounting.

NOTES

[1] John Carroll (1987), pp. 255–6.

[2] *Ibid.*, p. 249.

[3] Carroll claims on p. 251 that "$\mathbf{p}^*(t)$ could be interpreted as the probability that game [t] will be played," but we take it that it should read '$\mathbf{p}^*(t-1)$', given his definition of it in terms of $\mathbf{p}(t)$.

[4] See John Harsanyi (1977), p. 33, for a discussion of utility functions and the expected utility property.

[5] See David Kreps *et al.* (1982), pp. 245–252. They show that there can be cooperative equilibria when there is not common knowledge of the players' options or motivation.

[6] Carroll, *op. cit.*, p. 250.

[7] Drew Fudenberg and Eric Maskin (1986), especially p. 537.

REFERENCES

Carroll, J. W.: 1987, 'Indefinite Terminating Points and the Iterated Prisoner's Dilemma', *Theory and Decision* **22**, 247–256.

Fudenberg, Drew and Maskin, Eric: 1986, 'The Folk Theorem in Repeated Games with Discounting or with Incomplete Information', *Econometrica* **54**, 533–554.

Harsanyi, J. C.: 1977, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, Cambridge.

Kreps, D. M., Milgrom, P., Roberts, J., and Wilson, R.: 1982, 'Rational Cooperation in the Finitely Repeated Prisoner's Dilemma', *Journal of Economic Theory* **27**, 245–252.

Neal C. Becker
Department of Economics
University of Kansas
Lawrence, Ks. 66045, U.S.A.

and

Ann E. Cudd
Department of Philosophy
University of Kansas
Lawrence, Ks. 66045, U.S.A.