Rational Cooperation in the Finitely Repeated Prisoners' Dilemma

DAVID M. KREPS

Graduate School of Business, Stanford University, Stanford, California 94305

PAUL MILGROM

Graduate School of Management, Northwestern University, Evanston, Illinois 60201

JOHN ROBERTS AND ROBERT WILSON

Graduate School of Business, Stanford University, Stanford, California 94305

A common observation in experiments involving finite repetition of the prisoners' dilemma is that players do not always play the single-period dominant strategies ("finking"), but instead achieve some measure of cooperation. Yet finking at each stage is the only Nash equilibrium in the finitely repeated game. We show here how incomplete information about one or both players' options, motivation or behavior can explain the observed cooperation. Specifically, we provide a bound on the number of rounds at which Fink may be played, when one player may possibly be committed to a "Tit-for-Tat" strategy. *Journal of Economic Literature* Classification Numbers: 026, 213.

The purpose of this note is to demonstrate how reputation effects due to informational asymmetries can generate cooperative behavior in finitely repeated versions of the classic prisoners' dilemma. The methods employed are those developed in our work on the chain-store paradox (Kreps and Wilson [2], Milgrom and Roberts [4]). We refer the reader to those papers for motivation, formal definitions, and interpretation.

The basic game that we consider consists of N repetitions of the following two person, bimatrix, stage game:

	COL	
ROW	Fink	Cooperate
Fink	0, 0	a, b
Cooperate	b, a	1, 1

We require $a > 1$, $b < 0$, and $a + b < 2$.[1] At each stage, each of the two players, ROW and COL, recalls his previous actions and is informed about those of his opponent. The players move simultaneously at each stage. Payoffs in the overall game are the (undiscounted) sums of the stage payoffs.

This game has a unique Nash equilibrium path, which involves each player choosing to fink at every stage. The logic is similar to Selten's backwards induction in the chain–store game (although the argument there shows the uniqueness of the perfect equilibrium). In the final stage (which we call stage 1), finking strongly dominates cooperating, and so must ensue. Then, in the penultimate stage, finking does better than cooperating in terms of the current stage, while the choice at this stage cannot affect the outcome in stage 1. Thus finking will again be adopted by both players. And so on, for any finite N.[2] This outcome is clearly and dramatically inefficient.

This uniqueness result is disturbing in light of experiments with this game, of which there have been a very large number. (See Axelrod [1] and Smale [5] for references.) A common pattern in these experiments is that, at least for some time, both players cooperate and, in the process, end up with payoffs that are strictly greater than they would obtain under equilibrium play. The issue then is whether this puzzle can be resolved in the context of rational, self-interested behavior. The approach we adopt is to admit a "small amount" of the "right kind" of incomplete information.

In fact, we are able to show that certain kinds of informational asymmetries *must* yield a significant measure of cooperation in equilibrium, and that other plausible asymmetries may produce cooperation as well. Throughout, the equilibrium concept is that of sequential equilibrium (Kreps and Wilson [3]). Sequential equilibrium in a game of incomplete information requires that the action taken by any player at any point in the game tree must be part of an optimal strategy from that point forward, given his beliefs about the evolution of the game to this point (which must, to the extent possible, be consistent with Bayesian updating on the hypothesis that the equilibrium strategies have been used to date) and given that future play will be governed by the equilibrium strategies. The various models we use parallel those in [2, 4]. Each involves some element of uncertainty in the mind of (at least) one player about the other, and they can all be viewed in terms of a lack of common knowledge (between ROW and COL) that both

---

[1] If $a + b > 2$, then the strategy of both cooperating at each stage is Pareto-dominated by alternating between fink–cooperate and cooperate–fink. Much of our analysis can be adapted to handle this case.

[2] Note the sharp contrast with the infinitely repeated case, where any average payoff vector in the intersection of the positive orthant and the convex hull of the four possible stage payoff vectors can be achieved through a perfect equilibrium. Note also that in the finitely repeated case, Nash equilibrium behavior *off* the equilibrium path may involve some cooperation. But finking is required *everywhere* in any perfect equilibrium.

are rational players playing precisely the game specified above. The possibilities for more detailed analysis of this model and its application in economic, political, and military contexts appear to be very rich. Various combinations of the authors hope to report on such work in the future.

Model 1: ROW Might Play Tit-for-Tat

The first approach we consider supposes that, when the game begins, one of the players (say COL) is not absolutely certain that the other (ROW) will play "rationally" according to the payoffs specified above. Specifically, COL assigns probability $1 - \delta$ to the possibility of a "rational" opponent, and he allows a (very small) chance, δ, that ROW has available only the Tit-for-Tat strategy.[3] The Tit-for-Tat strategy requires the player using it to begin by cooperating and then to cooperate at stage $n - 1$ if and only if his opponent cooperated at the preceding stage, n. It is worth noting that this strikingly simple and quite natural strategy emerged as the winner in Axelrod's prisoners' dilemma tournament [1].

To present a sequential equilibrium in full detail for this game is difficult. There is no question that such equilibria exist: See Kreps and Wilson [3, Prop. 1]. But the "end play" of such equilibria are very complex. So we shall be content here to prove that in any sequential equilibrium, the number of stages where one player or the other finks is bounded above by a constant depending on δ but independent of N. Further, if we restrict attention to sequential equilibria that are not Pareto-dominated by any other sequential equilibria, then there is cooperation in all but the last "few" stages.

We prove these statements in a number of steps. The statement of each step except the last should be prefaced: In every sequential equilibrium...

Step 1. ...if it becomes common knowledge[4] before some stage that ROW is rational, then both ROW and COL fink at this and every succeeding stage, and their payoffs from the remainder of the game are zero.

The proof is by induction on the number of stages remaining. It is apparent if there is only one stage remaining. Suppose that it is true if there

---

[3] An alternative way to model this is to assume that ROW has available *all* the strategies above, but that with probability δ, ROW's payoffs are not as above but rather make playing Tit-for-Tat strongly dominant. The results given below can be proved for this alternative model, although the simple "common knowledge" arguments that we use are no longer available, and slightly more complex arguments are required. An advantage of this alternative model is that it eases interpretation of the probability assessed by COL that ROW is the Tit-for-Tat player as ROW's "reputation."

[4] It is common knowledge that ROW is rational if both players know this, both know that both know this, *ad infinitum*. More formally, an event E is common knowledge between two individuals at a state $\omega \in \Omega$ if there is some A in the finest common coarsening (meet) of their information partitions with $\omega \in A \subseteq E$. The crucial role of common knowledge will be illustrated shortly.

are $n - 1$ or fewer stages to go. Then with n stages remaining, the rational ROW must foresee that his present choice of action cannot influence the future course of the game, since it will remain common knowledge that he is rational when stage $n - 1$ arrives. Therefore he will maximize his immediate payoff, which means finking. Similarly, COL anticipates that no matter what he does at this stage, finking will occur at all later stages. In this round, finking is strictly better, so COL finks as well. Since both sides fink, their payoffs are each zero, and the induction is complete.

Step 2. ...if COL finks at stage $n + 1$, then ROW finks at stage n with probability one.

If ROW did cooperate in these circumstances, it would become common knowledge that he was rational. (The "Tit-for-Tat" ROW does not have this action available.) Thus cooperation nets zero in the continuation game. But finking can do no worse than zero in the continuation game and it is strictly dominant in the stage game. Thus finking does strictly better overall. This means that ROW must fink with probability one.

Step 3. ...starting from any point in the game tree (i), where COL assesses probability q that ROW is the Tit-for-Tat player, (ii), where there are n stages to go, and (iii), where COL *cooperated* on the previous stage, the expected payoff to COL for the remainder of the game is at least $qn + b$.

To show this, consider the strategy for COL of cooperating until the next time that ROW finks, and then finking ever after. Against the Tit-for-Tat player, this yields a payoff of n. Against the rational ROW, it yields no worse than b. Thus it yields an expected payoff that is at least $qn + (1 - q) b \geqslant qn + b$, and any equilibrium strategy must do at least as well.

Step 4. ...starting from any point in the game tree (i), where COL assesses probability q that ROW is the Tit-for-Tat player, (ii), where there are n stages to go, and (iii), where COL *finked* on the previous stage, the expected payoff to COL for the remainder of the game is at least $q(n - 1) + 2b$.

Because ROW is sure to fink (see step 2), COL knows that his assessment in the subsequent stage will again be q. So by cooperating at this stage, COL gets b immediately and at least $q(n - 1) + b$ in the continuation game. His overall expected payoff can be no worse than the sum of these, or $q(n - 1) + 2b$.

Step 5. ...starting at a point in the game tree (i), where COL assesses probability q that ROW is the Tit-for-Tat player, and (ii), where there are n stages to go, the expected payoff to the rational ROW player is no less than $q(n - 1) + 3b - a$.

Note first that COL will do no worse if the rational ROW plays Tit-for-Tat than if the rational ROW plays his equilibrium strategy. This is easily verified inductively, using steps (1) and (2). Thus the bounds obtained in steps (3) and (4) apply equally well if the rational ROW were to play Tit-for-Tat. And by playing Tit-for-Tat, the rational ROW nets within $b - a$ of whatever COL gets, path by path. This gives us the bound on ROW's payoff stated above.

Step 6.   ...if COL assesses probability q that ROW is the Tit-for-Tat player, and if there are more than $(2a - 4b + 2q)/q$ stages left to go, then ROW plays the Tit-for-Tat strategy with probability one. Thus along the equilibrium path, until the first stage less than $(2a - 4b + 2\delta)/\delta$, COL infers nothing from the observed behavior of ROW, and COL's assessment that ROW is the Tit-for-Tat player remains at δ.

In light of Step 2, all that is needed here is to show that ROW cooperates if COL has just cooperated in these circumstances. (The second part of the statement follows trivially from the first.) If ROW were to fink, it would become common knowledge that ROW is rational. Thus the total payoff from finking cannot exceed a—ROW gets at most a immediately (if COL cooperates) and then zero in the continuation game (by Step 1)). By cooperating, ROW will do no worse than b in this round (if COL finks) and, by Step 5, $q(n - 2) + 3b - a$ in the continuation game, where n is the number of stages remaining. If n exceeds $(2a - 4b + 2q)/q$, then cooperating is strictly better.

Step 7.   ...the total number of stages where one side or the other finks is bounded above by

$$\frac{2a - 4b + 2\delta}{\delta} \cdot \left[1 + \frac{2}{\min\{2 - a - b, 1\}} \right].$$

As seen in Step 6, ROW plays Tit-for-Tat until stage $(2a - 4b + 2\delta)/\delta$. If COL cooperates until ROW finks and then finks thereafter, his payoff must be at least $N - (2a - 4b + 2\delta)/\delta$. If COL finks before this date, then in that stage he gets a. If he then returns m stages later to cooperating, he gets b in the stage where he cooperates and zero in between. Thus he gives up $1 + m - a - b$ in this circumstance. A string of finks costs him $1 + (1 - (a + b))/m$ per round in comparison to cooperating. Thus, each time COL finks it costs him at least $\min\{2 - a - b, 1\}$. If he finks k times prior to stage $(2a - 4b + 2\delta)/\delta$, his payoff cannot exceed $N - k \cdot \min\{2 - a - b, 1\}$. These two bounds on COL's payoffs yield $k \leqslant (2a - 4b + 2\delta)/(\delta \cdot \min\{2 - a - b, 1\})$. Each such act of finking by COL provokes a Tit-for-Tat response from ROW in the next round, so there are at most $2k$ rounds before stage $(2a - 4b + 2\delta)/\delta$ when finking occurs. Thus the maximum number of rounds with finking is that given above.

Step 8. In any sequential equilibrium that is not Pareto-dominated by some other sequential equilibrium, there is no finking along the equilibrium path when more than $1 + (2a - 4b + 2\delta)/\delta$ stages remain.

For any equilibrium where there is finking before this date, a Pareto-superior equilibrium consists of not having that finking, and then continuing to play the game as if it had occurred.

Note that these bounds are not tight: If $\delta = 1$ they yield $n = 10$ in Step 8 for $a = 1.5$ and $b = -1$, yet in this circumstance one should see finking only in the last period. The looseness of these bounds suggests the need for further work.

The Tit-for-Tat theme can also be developed so as to further emphasize the role of lack of common knowledge. This development is in the spirit of Milgrom and Roberts [4, Appendix B].

Suppose that there are three states of the world. In state 1, ROW is the Tit-for-Tat player; in stages 2 and 3 he is rational. ROW learns whether he is Tit-for-Tat or not—his information partition (at the outset) is $\{1\}$, $\{2, 3\}$. COL, on the other hand, is given the information partition $\{1, 2\}$, $\{3\}$: In state 3 he knows whether ROW is the Tit-for-Tat player; in state 2 he does not. Suppose that state 3 prevails with very high probability. Then with very high probability, Row is not Tit-for-Tat, and COL *knows* that ROW is not Tit-for-Tat. But ROW is not sure that COL knows this, and one can show that the qualitative results proved for model 1 hold here. ROW will play Tit-for-Tat until near the end of the game, hoping that COL will be "deceived." And COL will pretend to be "deceived" even if he is not, as this improves his lot as well.

Or consider a four state case. In state 1 ROW is the Tit-for-Tat player—in states 2, 3 and 4 he is not. ROW is endowed with the information partition $\{1\}$, $\{2, 3\}$, $\{4\}$; COL with $\{1, 2\}$, $\{3, 4\}$. State 4 prevails with probability close to one. Then with probability close to one, ROW is not Tit-for-Tat, COL knows that this is so, ROW knows that COL knows this, but COL is not sure that ROW knows that COL knows. Once more the qualitative results for the original model hold up—ROW tries to "deceive" COL, knowing full well that COL will not be deceived but will act as if he is, and COL will do this in the hope that ROW may be unaware that is COL is not being deceived. One could go on like this forever: The general structure is that ROW's information partition should involve sets $\{1\}$, $\{2, 3\}$,..., $\{2m, 2m + 1\}$,... and COL's should involve $\{1, 2\}$, $\{3, 4\}$,..., $\{2m + 1, 2m + 2\}$,... (with termination eventually). The point is simply that so long as it is not common knowledge that ROW is not Tit-for-Tat, cooperation until near the end of the game will be rational.

Model 2: *Two-Sided Uncertainty about the Stage Payoffs*

In Model 1, COL entertains a hypothesis about ROW's behavior that cannot be generated if ROW is rational and has some stage game payoffs that he sums to arrive at his overall payoff. That is, COL's hypothesis, in terms of ROW's "true" utility function, necessarily involves payoffs for ROW that cut across stages. We might then wonder: Can long-run cooperation be attained if the only alternative hypotheses that are allowed (besides the hypothesis that the player is rational with the given stage payoffs) involve changes in the stage game payoffs? (This approach is used in [2].) The answer is a qualified yes.

Suppose that *each* player originally assesses a small probability that his opponent "enjoys" cooperation when it is met by cooperation. Given our zero-one normalization, we model this by assuming that COL assigns a small probability $\delta > 0$ that $a < 1$ for ROW, and ROW entertains a similar hypothesis about COL. We can then produce a sequential equilibrium wherein each side cooperates until the last few stages of the game, although again the end-game play is rather complex. In this equilibrium, if either side ever fails to cooperate, then the other side takes this as a sure sign that the defector has stage game payoffs with $a > 1$, and the noncooperative equilibrium ensues. As the details of this equilibrium are quite complex, we refrain from giving them here. Note, however, that if we move directly to a continuous-time formulation of this game, as in Kreps and Wilson [2, Sect. 4], then one equilibrium has cooperation throughout.

There are two qualifications to be made. First, two-sided uncertainty is required. If ROW, say, is uncertain about COL's stage payoffs, but it is common knowledge that $a > 1$ for ROW, then the only sequential equilibrium has finking throughout. (This is true for any "incomplete information" about one player's stage payoffs.) The second qualification, and certainly the more important, is that this game admits sequential equilibria in which long-run cooperation does not ensue, unlike the game with a Tit-for-Tat possibility. This is true even if we make a "plausibility" restriction on beliefs off the equilibrium path in the spirit of Section 3 of Kreps and Wilson [2]. Cooperation here requires a "boot-strapping" operation: Even if each side is *certain* that the other has $a < 1$, cooperation ensues only if each side hypothesizes that the other side will cooperate. (This is a fancy way of saying: If both sides have payoffs with $a < 1$, then there are two Nash equilibria in the stage game.) One might justify the cooperative equilibrium on "efficiency" grounds, but one cannot guarantee that cooperation will prevail in every sequential equilibrium.

ACKNOWLEDGMENTS

REFERENCES

1. R. AXELROD, The emergence of cooperation among egoists, *Amer. Pol. Sci. Rev.* **75** (1981), 306–318.
2. D. KREPS AND R. WILSON, Reputation and imperfect information, *J. Econ. Theory* **27** (1982), 253–279.
3. D. KREPS AND R. WILSON, Sequential equilibrium, *Econometrica*, in press.
4. P. MILGROM AND J. ROBERTS, Predation, reputation and entry deterrence, *J. Econ. Theory* **27** (1982), 280–312.
5. S. SMALE, The prisoner's dilemma and dynamical systems associated to non-cooperative games, *Econometrica* **48** (1980), 1617–1634.