

Cooperation in the Short and in the Long Run

H. PEYTON YOUNG

School of Public Affairs, University of Maryland, College Park, Maryland 20742

AND

DEAN FOSTER

Graduate School of Business, University of Chicago, Chicago, Illinois 60637

Received October 18, 1990

The long-run behavior of economic and biological processes is often dramatically altered when stochastic influences are taken into account. In fact, the smaller the noise, the more drastic the change can be. This seemingly paradoxical point is illustrated with the evolution of cooperation in repeated Prisoner's Dilemma. *Journal of Economic Literature* Classification Numbers: 026. © 1991 Academic Press, Inc.

The emergence of cooperative behavior in a competitive world poses something of a puzzle for classical theories of competition, since it appears to be inconsistent with the pursuit of self-interest by individuals. Of course, cooperation among closely related individuals makes sense if it increases the fitness of a specific gene (Hamilton, 1964). But this leaves open the question of why cooperation is often observed among *unrelated* individuals when "cheating" would yield a higher payoff for any one of them. An ingenious game-theoretic explanation for this case has been put forward by Axelrod and Hamilton (1981) and Axelrod (1984). Imagine a large population of individuals who engage in pairwise interactions. Every time that two individuals meet, they play a "game" whose outcome affects the number of offspring that each of them leaves in the next period. The fittest strategy is the one with the highest reproductive success rate. Even though a strategy is temporarily successful, however, it may eventually become less fit as the frequency of the other strategies in the population changes. In particular, a necessary condition for a strategy's con-

tinued success is that it do well *against itself* in addition to doing well against others. Reciprocal cooperation has this property, whereas purely noncooperative behavior does not.

Axelrod and Hamilton demonstrated this point in a celebrated series of experiments on the Prisoner's Dilemma (PD) game that will be described below (Axelrod and Hamilton, 1981; Axelrod, 1984). Their results have led to a good deal of optimism about the long-run viability of cooperation in a competitive environment. We shall argue, however, that the supposed advantage of *tit-for-tat* does not necessarily hold when their model is made more realistic. In particular, if the payoff rates are assumed to be slightly *variable*, then *tit-for-tat* is favored in the short run but not necessarily in the long run.

Our primary purpose, however, is not to argue for or against the viability of *tit-for-tat* in the Prisoner's Dilemma game. Indeed our results show that the verdict is quite sensitive to the precise way in which the model is specified. Rather, our goal is to call attention to an often-neglected property of evolutionary processes. When stochastic effects are incorporated into a dynamical system, not only can it alter its long-run behavior, but, paradoxically, the alteration may become *larger* as the noise level becomes *smaller* (Freidlin and Wentzell, 1984; Foster and Young, 1990). This feature of stochastic dynamical systems has important implications for biological and economic models of competition, which typically describe the macro-behavior of a system on the basis of many interactions among individuals. The usual assumption in these models is that the populations are large; hence, whatever variability may exist in the outcome of individual interactions tends to be smoothed out in the aggregate. It therefore seems reasonable to use expectations, since the stochastic variation is small. Unfortunately, this simplification can give quite misleading results, as we shall demonstrate. We choose Prisoner's Dilemma to illustrate this point for three reasons. First, the game is quite simple and is familiar to many readers. Second, there is a natural source of noise in the model, namely, the number of rounds of play in which each pair of individuals engages. Third, the stochastic and deterministic versions of the evolutionary model lead to strikingly different results.

THE EVOLUTIONARY MODEL OF REPEATED PRISONER'S DILEMMA

The payoffs from the one-shot Prisoner's Dilemma game are as follows:

	Cooperate	Defect
Cooperate	Each player gets 3	Cooperator gets 0 Defector gets 5
Defect	Cooperator gets 0 Defector gets 5	Each player gets 1

In the iterated version of this game (abbreviated IPD), two players engage in a series of one-shot PD games. Both players know that the engagement will end after the current round with some stated probability s called the *stopping probability*. In other words, the first round occurs for sure, the second round occurs with probability $1 - s$, and so forth. The expected number of rounds per engagement is $1/s$.

Axelrod's evolutionary experiment was conducted in the following manner. Contestants were invited to submit strategies for playing IPD, where the stopping probability was stated to be $s = .00346$. Several hundred strategies were entered in the tournament. Every entry was paired against every other entry (including itself) for five engagements of IPD. The lengths of the five engagements were 63, 77, 151, 156, and 308, which was the result of a single random draw from the distribution associated with the stopping probability .00346. (The results of this draw were not known to the contestants of course.) This tournament produced a total score for each strategy. At time zero the "population" consisted of the strategies originally entered into the contest. In each subsequent period, the relative frequency of each strategy in the population was determined by multiplying its frequency in the previous period by its tournament score. The scores therefore played the role of reproductive rates. Note that this evolutionary process is completely deterministic, for once the initial distribution of strategies is specified, then so is the distribution in all subsequent periods. If the probabilistic stopping rule were adopted, however, then the payoffs in each period would be a random variable.

To illustrate how this apparently minor modification can change the long-run behavior of the system in a major way, we shall consider a tournament with just three strategies: always cooperate (C), always defect (D), and tit-for-tat (T). At each time t , let $\mathbf{n}(t) = (n_C(t), n_D(t), n_T(t))$ denote the number of C-, D-, and T-players in the population, respectively, and let $N(t)$ be the total number of individuals. Similarly, let $\mathbf{p}(t) = (p_C(t), p_D(t), p_T(t)) = \mathbf{n}(t)/N(t)$ denote the proportions of C-, D-, and T-players in the population.

BEHAVIOR OF THE DETERMINISTIC SYSTEM WITH EXPECTED VALUES

For the moment, let us ignore the stochastic component and examine the behavior of the evolutionary process when expected values are used. Suppose that a D-player meets a T-player. In round 1 of the engagement, D's payoff is 5 and T's is 0. Thereafter, both get 1 in each round. The expected number of rounds after the first is $1/s - 1$. Therefore D's expected payoff is $5 + (1/s - 1) = 4 + 1/s$. In this manner one derives the following expected payoff matrix:

$$\mathbf{A} = \begin{pmatrix} 3/s & 0 & 3/s \\ 5/s & 1/s & 4 + 1/s \\ 3/s & 1/s - 1 & 3/s \end{pmatrix}.$$

The populations evolve over time according to the discrete-time dynamical equation

$$n_i(t + 1) = n_i(t) \mathbf{A}_i \mathbf{n}(t),$$

where $i = C, D, T$ and \mathbf{A}_i denotes the i th row of \mathbf{A} . The population proportions evolve according to the equation

$$p_i(t + 1) = p_i(t) \mathbf{A}_i \mathbf{p}(t) / [\mathbf{p}(t) \mathbf{A} \mathbf{p}(t)]. \quad (1)$$

The state space of this dynamical system is diagrammed in Fig. 1 for a stopping probability of .20. Each point in the triangle represents a set of population proportions. For example, the point D represents a population consisting entirely of defectors. The midpoint of the line DT represents half defectors, half T-players, and no cooperators, etc. At each point, the arrow indicates the instantaneous direction in which the proportions evolve. Beginning at any initial point such as P , the system evolves along a predictable, evolutionary path (shown here as a dotted line). Any point to the right of the line SS' evolves toward T, and any point to the left of the line evolves toward D. If the T-players have enough of an initial

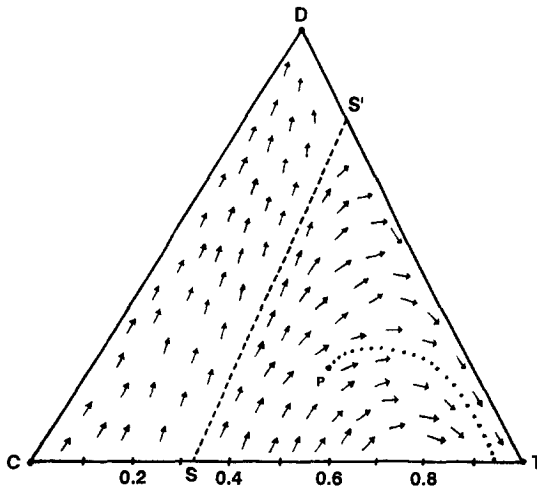


FIG. 1. The dynamical system with $s = .20$.

foothold relative to the D-players, then the former will be evolutionarily favored, otherwise the latter will be favored.

The strategy D is *locally stable* in the sense that any path beginning in a sufficiently small neighborhood of D converges to D. This notion may be interpreted in an evolutionary context as follows. Suppose that at some point in time the system consists only of defectors. Next, suppose that, by immigration or mutation, a small number of newcomers arrive on the scene playing C- or T-strategies. Then the system is pushed a little bit away from the point D. Since D is locally stable, however, the forces of selection will eventually cause the system to revert to D. In the language of biology, the strategy D is *evolutionarily stable* (Maynard Smith, 1982).

At first glance, it would appear that T is also evolutionarily stable, but this is not quite true. The problem is that once the system is pushed away from T, then the process will evolve back *in the general direction of* T but may not reach it. This will happen if the D-players die out before the C-players do. In this case, the evolutionary path ends somewhere near T on the CT line, as we have shown in Fig. 1. Any point on the CT line is stationary, because C-players and T-players are equally fit when there are no D-players to distinguish between them. (The fact that tit-for-tat is not as ESS was originally pointed out by Selten and Hammerstein (1984).)

NONEXTINCTION

It might be objected that T is disqualified from being evolutionarily stable on a mere technicality. And in a certain sense this is true. The failure of T to be evolutionarily stable depends on the assumption that strategies can become permanently extinct. If we take the biological metaphor seriously, however, then this is not a very natural assumption *in the long run*. First, most populations do not exist in isolation, but are subject to occasional immigration from outside. Second, background mutation is constantly at work to reinject new (and previously existing) strategies into the system. In this context, extinction is an unlikely event. Moreover, if mutation can reinstate anything that existed before, then *permanent extinction* is a zero-probability event. It seems reasonable, therefore, to examine the behavior of such systems when no strategy is allowed to die out completely. Specifically, we shall assume that mutation injects a small number of new C-players, T-players, and D-players into the population in each period, and that the number of mutants of each type is proportional to the population size (i.e., the rate of mutation is constant). Effectively, then, the evolutionary process stays bounded away from the edges, as shown in Fig. 2. In this case the point T*—which represents the situation in which all but newcomers are playing T—is a *bonafide* point attractor

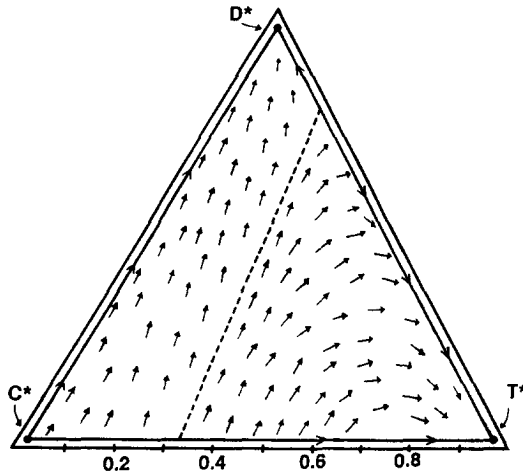


FIG. 2. The dynamical system with a small background mutation rate that keeps it away from the boundary.

toward which the system evolves whenever there are enough T-players to start with.

BEHAVIOR OF THE STOCHASTIC SYSTEM

Let us recall now that the length of play between two protagonists is actually a random variable determined by the stopping probability. For example, the payoff to a D-player in an engagement with a T-player is $5 + X$, where X is a geometric random variable with mean $(1 - s)/s$ and standard deviation $\sqrt{(1 - s)/s}$. Note that the longer the engagement is, the larger the payoff to D (and also to T). While this convention may seem somewhat unrealistic, the important point is that the payoffs between each pair of players are variable, which is realistic.

Suppose, for simplicity, that there is a fixed number of engagements in each time period, and that the number of engagements involving each pair of strategies is proportional to their frequencies in the population. Assume further that the interactions between different pairs of players are independent. Then the populations evolve according to a stochastic dynamical equation of the form

$$n_i(t + 1) = n_i(t) \hat{\mathbf{A}}_i \mathbf{n}(t) + rN(t), \quad (2)$$

where r is the mutation rate and $\hat{\mathbf{A}}$ is a random matrix whose expectation is \mathbf{A} . Each term of $\hat{\mathbf{A}}$ is the mean of independent, identically distributed random variables. If $N(t)$ is reasonably large, then the central limit theo-

rem implies that each term is approximately normally distributed with standard deviation proportional to $1/N(t)$.

We are interested in the behavior of the system when the population is kept fixed at some level N . We may think of N as the equilibrium size of the population relative to the rest of the environment due to the fact that individuals die out or migrate.

Consider first the case when N is small. Then the stochastic term is large, and the process bounces erratically around the state space. The basins of attraction therefore exert little influence, because the noise overwhelms the underlying selective drift in the system. Consider now the opposite case in which the population is large and the stochastic variability is small. In this situation, one would suppose that the system is likely to evolve along a path that is *close* to the path that it would have taken in the deterministic case. And indeed this is the case. In the short run, the behavior of the stochastic process is likely to be close to its behavior in the deterministic situation. There is a complication however. Once the process reaches a boundary—say the point T^* —it does not settle down. Small fluctuations continue to push it in various directions. Usually it returns toward T^* , but occasionally it is pushed quite far away—perhaps all the way into the D^* -basin. This is not a permanent disaster, though, because chance events will eventually push it into the T^* -basin again.

How often is the process in the D^* -basin as compared to the T^* -basin? One way of answering this question is by computer simulation. In each period, let the populations evolve according to the dynamical equation (2), but cull out excess members in proportion to their frequencies to keep the total number of individuals constant at level N . Assume a mutation rate of 0.1%; that is, new C-, D-, and T-players are added in each period at the rate of one per thousand of the existing population.

Beginning with equal numbers of C-, D-, and T-players, we ran the process for one million periods. Every 10 periods we recorded the current population frequencies. Figure 3 plots the results for a population of 100 individuals, where each dot represents one observation in the run, and there are a total of 100,000 observations. The simulation shows that the chances are about .59 that the population consists of at least 90% T-players, and about .18 that it consists of at least 90% D-players. (These percentages cannot be estimated easily from the figure, because the 59,000 points in the T corner are mostly overstruck.) In other words, tit-for-tat is favored in the sense that, 59% of the time, almost all of the population is composed of T-players. Note also that the process is *either* near D^* *or* near T^* with high probability. Intermediate regimes, or those involving many cooperators, are quite improbable.

But as N increases (and the noise decreases), the outcome is reversed. When N is 300, for example, the probabilities favor D instead of T (see

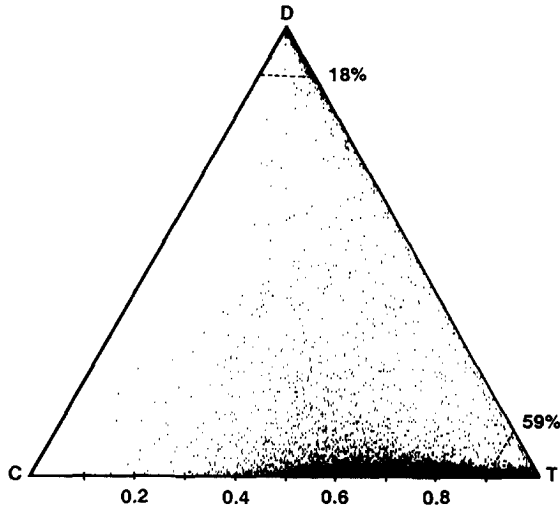


FIG. 3. Limiting distribution for $P = 100$.

Fig. 4). In this case the probability is 81% that most of the population is playing D. As N increases further, the probability becomes overwhelming that, in the very long run, almost everyone is a defector (see Fig. 5).

THE PROBLEM OF DEGENERATION

Perhaps even more interesting than this finding is the explanation for it. Consider the situation when the process is close to T^* . The defectors (and the cooperators) have almost all died out. In this relatively tranquil environment, pure cooperation has become almost as fit as tit-for-tat, because there are few defectors at large to punish the naivete of the cooperators. Moreover, there is a chance that the C-players will actually do better than the T-players for a certain length of time. This would occur, for example, if C-C engagements happened to be longer than T-T engagements. Of course, it is improbable that this bias would persist for very long, or that it would be large enough to negate the occasional C-D encounters. Nevertheless it could happen, and given a sufficiently long time horizon, it almost certainly will happen.

In other words, there is a small but positive probability that the process will start to creep along the line between C^* and T^* . At any moment, of course, chance events might favor the defectors, at which point the cooperators would be quickly weeded out and the process would circle back to T^* . There is, nevertheless, a small but positive probability that the weed-

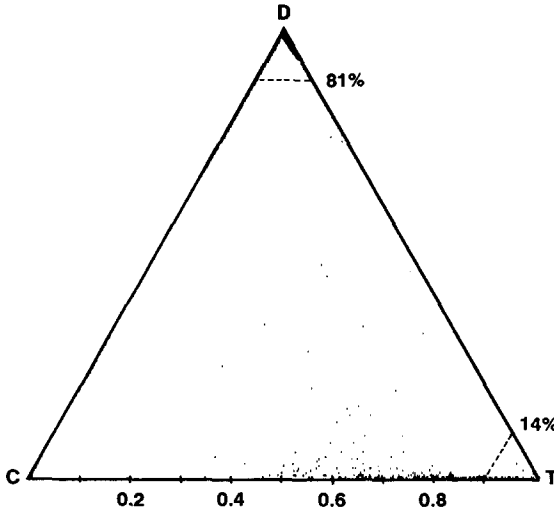


FIG. 4. Limiting distribution for $N = 300$.

ing out process will not occur until it is “too late,” that is, until the C-players have become dominant. In this event, once the defectors reassert themselves, the process will move rapidly toward the D*-corner, and both the C-players and the T-players will be almost wiped out.

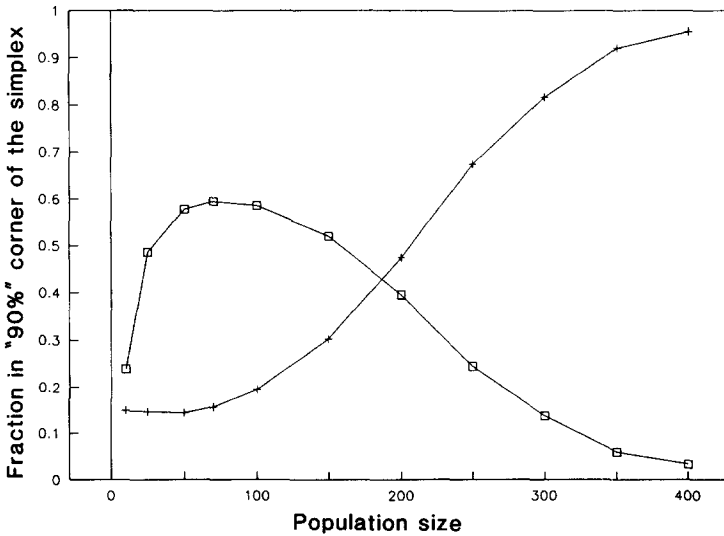


FIG. 5. Long-run probability that the population consists of 90% T-players (alternatively, 90% D-players) for various population sizes: (□) tit-for-tat, (+) defect.

This is still not the end of the story, however. Eventually, through chance events, the process will climb out of the D^* -basin and evolve once again toward T^* . But overcoming the resistance between D^* and the separatrix SS' by small chance nudges is extremely difficult and unlikely. Like Sisyphus condemned to roll the rock uphill, the process will make it partway up, only to lose headway and roll back down again. (Unlike Sisyphus, however, there is a positive probability that the process will eventually make it all the way up.) By contrast, the T^* -basin is very shallow along its lower boundary. This follows because cooperate and tit-for-tat are almost equally fit when the background level of defectors is very low. Thus there is very little uphill resistance in going from T^* to the separatrix, from which it is all downhill to D^* . In sum, climbing out of the D^* -basin in a given length of time is much more difficult (i.e., less likely) than creeping away from T^* on a nearly level field.

This intuitive argument can be made analytically rigorous using the theory of stochastic dynamical systems (Freidlin and Wentzell, 1984; Foster and Young, 1990). The basic idea is to compute the path of least resistance between the two locally stable regimes D^* and T^* . If mutations are sufficiently small, then the path of least resistance from T^* to D^* runs toward C^* , passes the separatrix, then drops down toward D^* . The resistance along this path is lower than the resistance along *any* path from D^* to T^* . From this it follows that the expected transition time from D^* to T^* becomes arbitrarily large relative to the expected transition time from T^* to D^* as the noise becomes small. Hence the process will be found in a neighborhood of D with near certainty as the population becomes large.

THE DEGENERATION OF NONCOOPERATIVE BEHAVIOR

Thus far we have focussed on the possibility that tit-for-tat might not be viable in the long run because it is undermined by less fit strategies. But of course the same arguments can be applied to noncooperative behavior. Defect will not be viable in the long run if it is undermined by strategies that only look like defect, that are cooperators in defectors' clothing. Consider, for example, the following strategy: defect on the first round and continue to defect until the opponent cooperates; thereafter play tit-for-tat. This will be called disguised tit-for-tat (DT). Suppose that the tournament consists initially of the three strategies D , T , and DT . Then, from any initial frequency distribution, the process will evolve toward tit-for-tat, assuming a small positive mutation rate and a small amount of noise. Thus defect may also be vulnerable to degeneration. In general, one cannot conclude very much about the long-run viability of any one

strategy without making a detailed study of the potential function associated with the whole system.

OTHER SOURCES OF VARIABILITY

The sole source of variability considered so far arises from the random length of engagements between different players. Many other sources of noise could also be considered. One possibility is noisy channels of communication. For example, a player might misread the action taken by his opponent with some small probability (Axelrod, 1984, p. 183; Molander, 1985; Donninger, 1986). More generally, one could simply postulate that the payoffs from any engagement between any two players are random variables with finite variances and expectations equal to the corresponding payoffs in the matrix A . If the number of engagements per period is large, and payoffs from different engagements are statistically independent, then (by the central limit theorem) the variability in the population *proportions* will be approximately normally distributed and the overall noise level will be inversely proportional to the population.

Yet another possibility is to assume, not that every player meets every other in a given period, but that a fixed proportion of the players are drawn at random to play in each period. If the population is large, then this variability is also well-approximated by a normal distribution, and the variance is proportional to the inverse of the population. The above arguments depend only on this assumption.

CONCLUSION

We close by noting that these arguments are just a special case of a well-known problem in evolutionary theory. If the selection pressure between two genes is very slight, then there is a chance that the less-fit gene will become fixed in the population due to random drift (Crow, 1986). In the first example the two genes are cooperate (the less fit) and tit-for-tat (the more fit). Unless there are a sufficient number of defectors at large to keep up selection pressure against cooperators, the cooperators may replace the T-players. To put it another way, if the environment is not sufficiently competitive for a period, then the genetic make up of an initially fit population may degenerate. It will then be decimated once competitive forces reassert themselves. Substantial variability—including a high background rate of mutation and immigration—is necessary to prevent such an outcome. For the evolution of cooperation, the moral is simply this: unless cooperative strategies are constantly being tested by

noncooperative ones, cooperation is viable in the short run, but not necessarily in the long run.

REFERENCES

- AXELROD R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- AXELROD R., AND HAMILTON, W. (1981). *Science* **211**, 1390.
- CROW, J. F. (1986). *Basic Concepts in Population, Quantitative, and Evolutionary Genetics*, Chap. 7. New York: Freeman.
- DONNINGER, C. (1986). in *Paradoxical Effects of Social Behavior*, (A. Diekmann and P. Mitter, Eds.) p. 123. Heidelberg: Physica-Verlag.
- FOSTER, D., AND YOUNG, H. P. (1990). *Theoret. Population Biol.* **38**, 219–232.
- FREIDLIN M. I., AND WENTZELL A. D. (1984). *Random Perturbations of Dynamical Systems*. New York: Springer-Verlag.
- HAMILTON, W. D. (1964). "The Genetical Evolution of Social Behavior, I and II," *J. Theoret. Biol.* **7**, 17–52.
- MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge Univ. Press.
- MOLANDER, P. (1985). *J. Conflict Resolution* **29**, 611.
- SELTEN, R., AND HAMMERSTEIN, P. (1984). *Behav. Brain Sci.* **7**(1), 115.