Tit-for-tat or win-stay, lose-shift?

Lorens A. Imhof[a,*], Drew Fudenberg[b], Martin A. Nowak[c,d,e]

[a]*Statistische Abteilung, Universität Bonn, D-53113 Bonn, Germany*
[b]*Department of Economics, Harvard University, Cambridge, MA 02138, USA*
[c]*Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA*
[d]*Department of Mathematics, Harvard University, Cambridge, MA 02138, USA*
[e]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA*

Abstract

The repeated Prisoner's Dilemma is usually known as a story of tit-for-tat (TFT). This remarkable strategy has won both of Robert Axelrod's tournaments. TFT does whatever the opponent has done in the previous round. It will cooperate if the opponent has cooperated, and it will defect if the opponent has defected. But TFT has two weaknesses: (i) it cannot correct mistakes (erroneous moves) and (ii) a population of TFT players is undermined by random drift when mutant strategies appear which play always-cooperate (ALLC). Another equally simple strategy called 'win-stay, lose-shift' (WSLS) has neither of these two disadvantages. WSLS repeats the previous move if the resulting payoff has met its aspiration level and changes otherwise. Here, we use a novel approach of stochastic evolutionary game dynamics in finite populations to study mutation–selection dynamics in the presence of erroneous moves. We compare four strategies: always-defect (ALLD), ALLC, TFT and WSLS. There are two possible outcomes: if the benefit of cooperation is below a critical value then ALLD is selected; if the benefit of cooperation is above this critical value then WSLS is selected. TFT is never selected in this evolutionary process, but lowers the selection threshold for WSLS.
© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

A crucial question of evolutionary biology is how natural selection can lead to cooperative behavior. A cooperator is someone who pays a cost, c, for another individual to receive a benefit, b. Cost and benefit are measured in terms of reproductive success. Various mechanisms for the evolution of cooperation have been proposed including kin selection (Hamilton, 1964), group selection (Wilson, 1975; Wynne-Edwards, 1962), indirect reciprocity (Nowak and Sigmund, 1998, 2005) and spatial (or network) reciprocity (Nowak and May, 1992; Ohtsuki et al., 2006; Skyrms, 2004). Here we analyze direct reciprocity which was introduced by Trivers (1971). The hallmark of direct reciprocity is the repeated Prisoner's Dilemma (PD) (Luce and Raiffa, 1957; Rapoport and Chammah, 1965).

In the PD, two players have a choice to cooperate, C, or to defect, D. If both cooperate they receive a payoff R. If both defect, they receive a lower payoff P. But a defector versus a cooperator gains the highest payoff, T, while the cooperator is left with the lowest payoff S. The game is a PD if $T > R > P > S$. In a non-repeated PD, it is best to defect. Cooperation becomes an option if the game is repeated. For the repeated PD, we also ask that $R > (T + S)/2$. Otherwise alternating between cooperation and defection would lead to a higher payoff than mutual cooperation.

There are many conceivable strategies for the repeated PD, and, in 1978, Axelrod decided to conduct computer tournaments in order to find successful strategies (Axelrod, 1984; Axelrod and Hamilton, 1981). There were 14 entries to the first tournament and 63 to the second. In both tournaments, the simplest of all strategies won. It was

*Corresponding author.
E-mail address: limhof@uni-bonn.de (L.A. Imhof).

tit-for-tat (TFT), submitted by Anatol Rapoport (see also Rapoport and Chammah, 1965). TFT cooperates in the first round and then does whatever the opponent has done in the previous round. It will play C for C and D for D. Hence, TFT embodies the simple principle 'an eye for an eye and a tooth for a tooth'. It was remarkable that TFT won both tournaments, because participants of the second tournament knew the outcome of the first and tried to design (more sophisticated) strategies that were superior to TFT. The success of TFT inspired much subsequent research (Dugatkin, 1988; Milinski, 1987; Wilkinson, 1984).

TFT, however, has a weakness which did not become apparent in the error-free environment of Axelrod's tournaments. If two TFT players interact with each other and one defects by mistake, then the other one will retaliate, which in turn leads to a sequence of alternating moves of cooperation and defection. In the presence of noise, in an infinitely repeated game, the expected payoff per round for TFT versus TFT is only $(R + S + T + P)/4$. TFT's unrelenting punishment never forgives even a single deviation, but forgiveness is needed to restore cooperation after occasional errors (Fudenberg and Maskin, 1990; Molander, 1985; Nowak and Sigmund, 1992).

Another drawback of TFT is the following. Imagine a mixed population of players using TFT and players using the strategy 'always-cooperate' (ALLC). In the absence of noise, both strategies have the same average payoff. The frequency (= relative abundance) of ALLC can increase by random drift. This means TFT is not evolutionarily stable (Boyd and Lorberbaum, 1987; Selten and Hammerstein, 1984). Once the abundance of ALLC has fluctuated above a certain threshold, players using 'always-defect' (ALLD) can invade the population.

Win-stay, lose-shift (WSLS) is a very simple strategy which has neither of these two weaknesses (Nowak and Sigmund, 1993). WSLS works as follows. If WSLS receives payoff T or R it will repeat the same move next round, this is 'win-stay'. If WSLS receives only payoff P or S, it will switch to the opposite move, this is 'lose-shift'. Thus, the choice of the next action is based on some aspiration level which lies between P and R. The problem of adapting the aspiration level has been considered e.g. in Posch et al. (1999). Another name for WSLS is 'perfect TFT'; this name reflects the fact that for typical payoffs the rule is a symmetric subgame-perfect equilibrium, while TFT is typically not. WSLS is also known as Pavlov (Kraines and Kraines, 1989, 1995). Milinski and Wedekind (1998) provide empirical evidence for WSLS. For longer memory versions of WSLS see Neill (2001) and Wakano and Yamamura (2001).

As illustrated in Fig. 1, WSLS allows (i) to correct mistakes and (ii) to exploit ALLC. Hence, a WSLS population receives a high average payoff even in the presence of errors and is stable against random drift to ALLC. However, WSLS is exploited by ALLD in every second round whereas TFT is not. Evolutionary analysis

Fig. 1. WSLS cooperates against itself most of the time: a mistake by one of the two players that use WSLS triggers a single round of punishment and then both return to cooperation. Moreover, WSLS can exploit unconditional cooperators, so that WSLS is proof against random drift to ALLC.

and computer simulations of heterogeneous populations using a set of probabilistic strategies indicate that WSLS outcompetes TFT in most cases (Nowak and Sigmund, 1993; Nowak et al., 1995). Fudenberg and Maskin (1990) considered a concept of evolutionary stability for repeated games with errors and showed that WSLS is, in this sense, evolutionarily stable in the set of strategies of finite complexity, while TFT is not.

2. Mutation–selection dynamics in the presence of erroneous moves

Here, we present an exact calculation that allows us to study the evolutionary dynamics of WSLS and TFT under mutation, selection and random drift in finite populations and in the presence of noise. We consider a population of N individuals. Each individual derives a payoff by interacting with (a representative ensemble of) all other individuals. The fitness of an individual is given by a constant, representing the baseline fitness, plus the payoff from the game. Weak selection means that the baseline fitness is large compared to the payoff. See Traulsen et al. (2006) for a discussion of the influence of the intensity of selection on evolutionary stability. In each time step, a random individual is chosen for reproduction proportional to fitness. The offspring replaces another randomly chosen individual. These dynamics describe a frequency-dependent Moran process (Nowak, 2006; Nowak et al., 2004), but we expect that similar processes, for example, those based on pairwise comparison and updating, would lead to qualitatively similar conclusions (Imhof and Nowak, 2006;

Traulsen et al., 2005, 2007). Reproduction can be genetic or cultural.

Mutation means that the offspring of one individual need not always adopt the parent's strategy, but with a small probability, u, a different, randomly chosen strategy is adopted (Kandori et al., 1993; Young, 1993); see Chapter 5 in Fudenberg and Levine (1998) for a survey of adjustment models with mutations. The resulting mutation–selection process leads to a stationary distribution, where each strategy is represented by a certain fraction of the population. We characterize this mutation–selection equilibrium in the limit of small mutation rate ($u \to 0$) and large population size ($N \to \infty$). Our approach is similar to that of Imhof et al. (2005), but here we consider the infinitely repeated PD with noise.

At first, we study a simplified version of the PD which is characterized only by two parameters, the cost, c, and the benefit, b, of cooperation. The payoff matrix is given by

$$
\begin{array}{c}
\quad C \quad\ \ D \\
\begin{array}{c} C \\ D \end{array}
\begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}.
\end{array}
\tag{1}
$$

This game is a PD if $b > c$. Note that the matrix in (1) describes a subset of possible PD games. The general case and all relevant proofs are discussed in the Appendix.

We begin by studying only two strategies, ALLD and TFT. Both strategies are best replies to themselves, but TFT has the larger basin of attraction. The equilibrium distribution of our mutation–selection process is entirely centered on TFT (Fig. 2a).

Let us now add ALLC and analyze the competition of three strategies: ALLD, TFT and ALLC. In this system, the equilibrium distribution is entirely centered on ALLD. Therefore, adding ALLC undermines the evolutionary success of TFT (Fig. 2b).

Next we consider ALLD and WSLS. The remarkably simple criterion

$$
b/c > 3
\tag{2}
$$

is a necessary condition for the selection of WSLS. If $b/c < 3$ then ALLD is selected. If $b/c > 3$ then WSLS is selected provided the intensity of selection is sufficiently weak. Adding ALLC does not change this outcome. In the three strategy system, ALLD, WSLS and ALLC, it is also the case that $b/c > 3$ allows the selection of WSLS for weak selection (Fig. 2c). For stronger selection, the critical threshold for b/c is somewhat larger than 3 (Fig. 3).

Now comes another surprise. Let us add TFT and therefore consider the four strategy system, ALLD, ALLC, WSLS and TFT. If

$$
b/c > 2
\tag{3}
$$

then WSLS is selected for any intensity of selection (Fig. 2d). If $b/c < 2$ then ALLD is selected. Remarkably, TFT is never selected, but the presence of TFT lowers the threshold for the evolution of cooperation. As noted previously (Nowak and Sigmund, 1992), TFT is a catalyst for the emergence of cooperation, but is not the target of evolution.

There is a nice intuitive justification for the critical thresholds in (2) and (3). In the absence of noise, the payoff
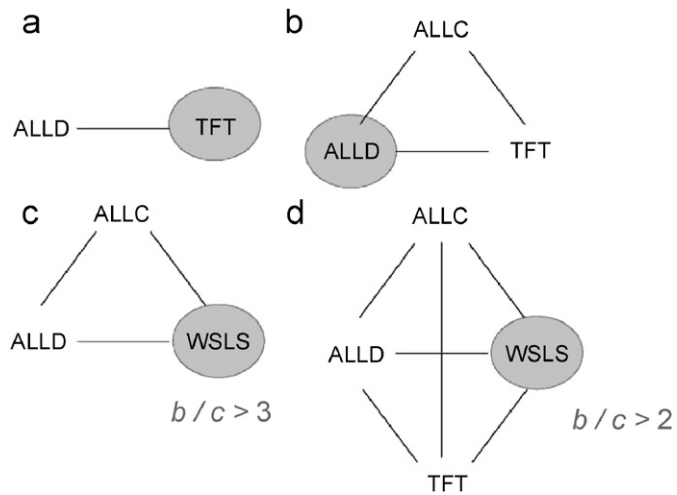

Fig. 2. Selected strategies in large populations when mutation rates are small. (a) ALLD against TFT is a coordination game, where TFT is always selected as long as the benefit, b, is larger than the cost, c. (b) If unconditional cooperators are added, they can invade and take over a TFT population by random drift, but an ALLC population is quickly taken over by ALLD players. ALLD is the selected strategy. (c) If instead of TFT, WSLS competes with ALLC and ALLD, then WSLS is selected, provided the benefit-to-cost ratio, b/c, is sufficiently large. The critical value is 3 in the limit of weak selection. (d) Similar behavior as in (c) is observed if TFT is added, but the critical value is now lowered to 2 for any intensity of selection.

Fig. 3. In the system with strategies ALLD and WSLS, there is a critical threshold for the benefit-to-cost ratio, b/c, that determines whether ALLD or WSLS is selected in large populations. WSLS is selected if and only if b/c exceeds the threshold. The figure shows how the critical value depends on the intensity of selection. The value is about 3.6 for strong selection, decreases with increasing background fitness and approaches 3 in the limit of weak selection. In particular, if $b/c > 3$, then WSLS is selected if the intensity of selection is sufficiently weak.

matrix for WSLS versus ALLD is given by

$$\begin{array}{c} \quad \text{WSLS} \quad \text{ALLD} \\ \begin{array}{c} \text{WSLS} \\ \text{ALLD} \end{array} \begin{pmatrix} b - c & -c/2 \\ b/2 & 0 \end{pmatrix}. \end{array} \quad (4)$$

The entries represent the average payoffs per round. Adding small amounts of noise introduces only small perturbations of the payoff values. Remember that WSLS cooperates against ALLD every other round. If we only compare WSLS and ALLD, then all that matters is which strategy is risk dominant (Fudenberg and Levine, 1998, p. 138; Nowak et al., 2004). WSLS is risk dominant if $b/c > 3$. If we add TFT, then an ALLD population is rapidly invaded by TFT (Imhof et al., 2005; Nowak et al., 2004) and the decisive criterion is that WSLS versus ALLD is a coordination game. This is the case if $b/c > 2$. If on the other hand $b/c < 2$, then WSLS is dominated by ALLD. This intuitive justification also explains the more complicated criteria for the general PD game, which are discussed in the Appendix.

We have focused on the four strategies, ALLC, ALLD, TFT and WSLS, because they represent important basic types of behavior in a repeated PD. It would be interesting to add further strategies, e.g. generous TFT (GTFT). GTFT cooperates if the opponent cooperated in the previous round and cooperates with a fixed probability $q > 0$ if the opponent defected. Thus GTFT can also correct mistakes although not as efficiently as WSLS. Unfortunately, adding GTFT makes it difficult to prove mathematical results for our stochastic model. The numerical results in Table 1 suggest that ALLD and GTFT may persist for a certain range of benefit–cost ratios but WSLS is again selected if that ratio is sufficiently large. Our findings hold for the simultaneous PD, but not for the alternating PD (Frean, 1994; Nowak and Sigmund, 1994), where WSLS is not error correcting. For the alternating PD one may consider strategies with longer memory that use ideas of WSLS and are able to correct mistakes (Neill, 2001). Further challenges for future research include finite population dynamics for the continuous PD (Killingback et al., 1999; Roberts and Sherratt, 1998; Wahl and Nowak, 1999a, 1999b), for general repeated games (Aumann and Shapley, 1976; Fudenberg and Maskin, 1986) and finally for repeated games with imperfectly observed actions (Fudenberg et al., 1994; Fudenberg and Levine, 1994).

In summary, we have studied stochastic game dynamics in populations of finite size under the influence of both mutation and selection. We consider the infinitely repeated PD with noise. Players make mistakes when implementing their strategies. In a population consisting of the three strategies, ALLC, ALLD and TFT, we find that the equilibrium distribution of our evolutionary process is entirely centered on ALLD. In contrast, a population consisting of ALLC, ALLD and WSLS settles predominantly on WSLS provided the benefit-to-cost ratio exceeds 3, $b/c > 3$. Otherwise the system chooses ALLD. Finally,

Table 1
Limit distributions of the stochastic system for the four strategies ALLC, ALLD, TFT and WSLS and of the augmented system which contains also GTFT

b	ALLC	ALLD	TFT	WSLS	
1.5	0.0005	0.9741	0.0179	0.0075	
2.0	0.0010	0.9201	0.0176	0.0613	
2.5	0.0001	0.0343	0.0006	0.9650	
3.0	0.0000	0.0000	0.0000	1.0000	
b	ALLC	ALLD	TFT	WSLS	GTFT, $q = \frac{1}{10}$
1.5	0.0001	0.3592	0.0036	0.0043	0.6328
2.0	0.0003	0.3028	0.0031	0.0201	0.6736
2.5	0.0001	0.0343	0.0003	0.9445	0.0207
3.0	0.0000	0.0000	0.0000	1.0000	0.0000
b	ALLC	ALLD	TFT	WSLS	GTFT, $q = \frac{1}{3}$
1.5	0.0003	0.9482	0.0092	0.0049	0.0375
2.0	0.0003	0.4570	0.0046	0.0503	0.4879
2.5	0.0000	0.0201	0.0002	0.9290	0.0506
3.0	0.0000	0.0000	0.0000	1.0000	0.0000
b	ALLC	ALLD	TFT	WSLS	GTFT, $q = \frac{1}{2}$
1.5	0.0003	0.9779	0.0095	0.0043	0.0080
2.0	0.0005	0.8912	0.0089	0.0385	0.0609
2.5	0.0000	0.0209	0.0002	0.9569	0.0220
3.0	0.0000	0.0000	0.0000	1.0000	0.0000

Here, GTFT cooperates with probability $q = 1/10$, $q = 1/3$ or $q = 1/2$, respectively, if the opponent defected in the previous round. The population size is $N = 10,000$, the cost of cooperation is $c = 1$ and the background fitness is $K = 10$.

we have studied a population consisting of four strategies, ALLC, ALLD, TFT and WSLS. If $b/c > 2$ then WSLS is selected, otherwise ALLD is selected. Note that TFT lowers the selection threshold for WSLS, but is never chosen itself. TFT is the catalyst for the emergence of cooperation, but WSLS is its target.

Appendix A. Evolutionary dynamics in finite populations

We consider a discrete-time frequency-dependent Moran process with mutations for a two-player game with n strategies and payoff matrix $A = (a_{ij})$. The possible states of a population of size N are given by the vectors $x = (x_1, \ldots, x_n) \in \mathbb{N}_0^n$ with $x_1 + \cdots + x_n = N$. Here x_i denotes the number of individuals using strategy i. If the baseline fitness is K and the current state of the population is x, then

the fitness of every individual that plays i is given by

$$f_i(x) = K + \frac{\sum_{j=1}^n a_{ij}x_j - a_{ii}}{N-1},$$

where we assume that individuals do not interact with themselves. At every time step, one individual is chosen for reproduction and the probability of being chosen is proportional to fitness. That is, the probability that an individual using strategy i reproduces is equal to $x_i f_i(x) / \sum_j x_j f_j(x)$. With probability $1-u$, the offspring uses the same strategy i. Otherwise, he mutates to one of the other $n-1$ strategies with each of them being equally likely. The offspring replaces a randomly chosen member of the current population. For positive u, this procedure gives rise to an irreducible Markov chain with a unique invariant distribution, denoted by $\pi = \pi(x; u, K, N)$.

Letting the mutation probability u go to zero, we obtain a limit distribution, which concentrates on the homogeneous states, where all individuals play the same strategy. Let $\pi_1^*(K, N), \ldots, \pi_n^*(K, N)$ denote the limits corresponding to strategies $1, \ldots, n$, respectively. In the case of 2×2 games, the frequency-dependent Moran process is a simple

small mutation rates u, the Moran process spends almost all its time at the homogeneous states, an amount of time of order $O(u)$ on the edges and only $o(u)$ at all other states.

We say that strategy i is selected, if $\lim_{N \to \infty} \pi_i^*(K, N) = 1$.

Appendix B. The repeated PD with errors

Consider the PD with payoffs $T > R > P > S > 0$ and suppose that $R > (T + S)/2$. We now derive the payoffs for the strategies ALLC, ALLD, TFT and WSLS in the infinitely repeated PD when in each round each player makes an error with probability $\varepsilon > 0$. The successive actions chosen by players using such strategies can be modeled by a Markov chain, see Nowak and Sigmund (1990). The presence of noise guarantees that the chain is ergodic. In particular, the initial action does not influence the average payoff per round, when infinitely many rounds are played. Computing the ergodic distribution of the realized actions for each pair of strategies, we obtain the payoff matrix in (B.1), where all terms of order $O(\varepsilon^2)$ have been ignored:

$$\begin{pmatrix} R & S & R & \frac{R+S}{2} \\ -\varepsilon(2R-S-T) & +\varepsilon(R+P-2S) & -\varepsilon(3R-T-2S) & +\frac{\varepsilon}{2}(T+P-R-S) \\[2ex] T & P & P & \frac{P+T}{2} \\ -\varepsilon(2T-R-P) & +\varepsilon(S+T-2P) & +\varepsilon(S+2T-3P) & +\frac{\varepsilon}{2}(R+S-T-P) \\[2ex] R & P & \frac{T+R+P+S}{4} & \frac{T+R+P+S}{4} \\ +\varepsilon(2T+S-3R) & +\varepsilon(T+2S-3P) & & \\[2ex] \frac{R+T}{2} & \frac{P+S}{2} & \frac{T+R+P+S}{4} & R \\ +\frac{\varepsilon}{2}(S+P-R-T) & +\frac{\varepsilon}{2}(R+T-S-P) & & +\varepsilon(T+2P+S-4R) \end{pmatrix}. \quad (B.1)$$

birth–death process, and using an explicit representation of the invariant distribution (Ewens, 2004), one may show that

$$\frac{\pi_1^*(K, N)}{\pi_2^*(K, N)} = \prod_{i=1}^{N-1} \frac{f_1(i, N-i)}{f_2(i, N-i)}. \quad (A.1)$$

For $n \times n$ games with $n \geqslant 3$, the limit distribution can be determined using the results of Fudenberg and Imhof (2006). This approach requires the computation of every fixation probability ρ_{ij}, which is the probability that in the absence of mutations, a single individual using strategy j takes over a population where otherwise only strategy i is used. In particular, one needs to consider only the no-mutation process along the edges, where just two different strategies are present. This stems from the fact that for

Appendix C. Selection results

1. *Competition of ALLD and TFT*: If $P > (T + R + P + S)/4$, then ALLD strictly dominates TFT. Here and throughout we assume that ε is sufficiently small. It follows from (A.1) that ALLD is selected in large populations. Suppose that $P < (T + R + P + S)/4$. Then ALLD against TFT is a coordination game. Approximating the logarithm of the right-hand side of (A.1) by Riemann sums, one may show that TFT is selected if and only if

$$\int_0^1 \log \frac{P + K + O(\varepsilon)}{[P + O(\varepsilon)]x + (T + R + P + S)(1-x)/4 + K} \, dx < 0.$$

But for ε small enough, the integral is obviously negative. In the simplified version of the PD which is characterized

by benefit b and cost c, we have $T + R + P + S = 2(b - c) > 0$ and $P = 0$, so that TFT is always selected.

2. *Competition of ALLC, ALLD and TFT*: It is shown in Example 2 of Fudenberg and Imhof (2006) that the limit distribution is given by $\pi^* = [\gamma_1 + \gamma_2 + \gamma_3]^{-1}(\gamma_1, \gamma_2, \gamma_3)$, where

$$\gamma_1 = \rho_{21}\rho_{31} + \rho_{21}\rho_{32} + \rho_{31}\rho_{23},$$
$$\gamma_2 = \rho_{31}\rho_{12} + \rho_{12}\rho_{32} + \rho_{32}\rho_{13},$$
$$\gamma_3 = \rho_{21}\rho_{13} + \rho_{12}\rho_{23} + \rho_{13}\rho_{23}.$$

Using a well-known explicit expression for the absorption probabilities (Karlin and Taylor, 1975) and an approximation based on Riemann sums, one may show that

$$\rho_{21}(N) = O(e^{-\beta N}), \quad \rho_{23}(N) = O(e^{-\beta N})$$

for some $\beta > 0$. On the other hand,

$$\inf_{N \geqslant 2} \rho_{12}(N) > 0, \quad \inf_{N \geqslant 2} \rho_{31}(N) > 0.$$

It is now obvious that $\pi_2^*(N) \to 1$, so that ALLD is selected.

3. *Competition of ALLD and WSLS*: If $(P + T)/2 > R$, then ALLD strictly dominates WSLS, and so, in view of (A.1), ALLD is selected. Suppose next that $(P + T)/2 < R$. Then ALLD against WSLS is a coordination game and which equilibrium is selected depends on the sign of

$$\gamma(K) = \int_0^1 \log \frac{(P + T)x + 2P(1 - x) + 2K}{2Rx + (P + S)(1 - x) + 2K} dx.$$

If $\gamma(K) > 0$, then ALLD is selected; if $\gamma(K) < 0$, WSLS is selected.

If the payoffs in the PD are given in terms of benefit b and cost c, and $b/c < 2$, then ALLD is strictly dominant, so that ALLD is selected. Suppose that $b/c > 2$. We have

$$\gamma(K) = \int_0^1 \log \frac{bx + 2K}{(2b - c)x - c + 2K} dx,$$

and it can be shown that

$$\gamma(K) < 0 \quad \text{for all } K \text{ sufficiently large} \iff \frac{b}{c} > 3.$$

More precisely, if $b/c > 3$, there is a critical value $K^*(b/c)$ such that if $K < K^*$, then $\gamma(K) > 0$, so that ALLD is selected, and if $K > K^*$, then $\gamma(K) < 0$, so that WSLS is selected. Fig. 3 shows the relation between K^* and the benefit–cost ratio. Notice that $K^* \to \infty$ when b/c approaches 3 from above. If $b/c < 3$, then $\gamma(K) > 0$ for all K, so that ALLD is selected independently of the intensity of selection.

4. *Competition of ALLC, ALLD and WSLS*: ALLD strictly dominates ALLC and WSLS strictly dominates ALLC. If $(P + T)/2 > R$, ALLD strictly dominates WSLS and ALLD is selected. Suppose next that $R > (P + T)/2$, so

that ALLD against WSLS is a coordination game. Let

$$\beta_{21} = \int_0^1 \log \frac{Tx + P(1 - x) + K}{Rx + S(1 - x) + K} dx,$$
$$\beta_{41} = \int_0^1 \log \frac{(R + T)x + 2R(1 - x) + 2K}{2Rx + (R + S)(1 - x) + 2K} dx,$$
$$\beta_{24} = \int_0^{z_{24}} \log \frac{(P + T)x + 2P(1 - x) + 2K}{2Rx + (P + S)(1 - x) + 2K} dx,$$
$$\beta_{42} = \int_0^{z_{42}} \log \frac{(P + S)x + 2R(1 - x) + 2K}{2Px + (P + T)(1 - x) + 2K} dx,$$

where

$$z_{24} = \frac{P - S}{2R - S - T}, \quad z_{42} = \frac{2R - P - T}{2R - S - T}.$$

If $\min\{\beta_{42}, \beta_{41}\} < \min\{\beta_{24}, \beta_{21}\}$, then ALLD is selected. If $\min\{\beta_{42}, \beta_{41}\} > \min\{\beta_{24}, \beta_{21}\}$, then WSLS is selected. Examining the integral conditions for $K \to \infty$, one may show that in the simplified version of the PD, WSLS is selected for K sufficiently large if and only if $b/c > 3$.

5. *Competition of ALLC, ALLD, TFT and WSLS*: If $R < (P + T)/2$, then ALLD is selected. Suppose $R > (P + T)/2$ and $P < (T + R + P + S)/4$. Then ALLD and WSLS form a coordination game and so do ALLD and TFT. In this case, WSLS is selected. Suppose next that $2R > P + T$ and $P > (T + R + P + S)/4$, so that again ALLD and WSLS form a coordination game, but ALLD strictly dominates TFT. In this case, either ALLD or WSLS is selected, and which one is selected can be determined by comparing certain integrals similar to those in the analysis for the competition between ALLC, ALLD and WSLS. This last case cannot occur in the simplified version of the PD characterized by benefit b and cost c, and it follows that ALLD is selected if $b/c < 2$, and WSLS is selected if $b/c > 2$.

References

Aumann, R., Shapley, L., 1976. Long term competition: a game theoretic analysis. Mimeo, Hebrew University.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York.

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.

Boyd, R., Lorberbaum, J.P., 1987. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. Nature 327, 58–59.

Dugatkin, L.A., 1988. Do guppies play tit for tat during predator inspection visits? Behav. Ecol. Sociobiol. 23, 395–397.

Ewens, W.J., 2004. Mathematical Population Genetics. Springer, New York.

Frean, M.R., 1994. The prisoner's dilemma without synchrony. Proc. R. Soc. London B 257, 75–79.

Fudenberg, D., Imhof, L.A., 2006. Imitation processes with small mutations. J. Econ. Theory 131, 251–262.

Fudenberg, D., Levine, D.K., 1994. Efficiency and observability with long-run and short-run players. J. Econ. Theory 62, 103–135.

Fudenberg, D., Levine, D.K., 1998. The Theory of Learning in Games. MIT Press, Cambridge, MA.

Fudenberg, D., Maskin, E., 1986. The folk theorem in repeated games with discounting or with incomplete information. Econometrica 54, 533–554.

Fudenberg, D., Maskin, E., 1990. Evolution and cooperation in noisy repeated games. Am. Econ. Rev. 80, 274–279.

Fudenberg, D., Levine, D., Maskin, E., 1994. The folk theorem with imperfect public information. Econometrica 62, 997–1039.

Hamilton, W.D., 1964. The genetical evolution of social behaviour. J. Theor. Biol. 7, 1–16.

Imhof, L.A., Nowak, M.A., 2006. Evolutionary game dynamics in a Wright–Fisher process. J. Math. Biol. 52, 667–681.

Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. Proc. Natl Acad. Sci. USA 102, 10797–10800.

Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29–56.

Karlin, S., Taylor, H.M., 1975. A First Course in Stochastic Processes. Academic Press, San Diego.

Killingback, T., Doebeli, M., Knowlton, N., 1999. Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. Proc. R. Soc. London B 266, 1723–1728.

Kraines, D., Kraines, V., 1989. Pavlov and the prisoner's dilemma. Theory Decision 26, 47–79.

Kraines, D., Kraines, V., 1995. Evolution of learning among Pavlov strategies in a competitive environment with noise. J. Conflict Resolut. 39, 439–466.

Luce, D., Raiffa, H., 1957. Games and Decisions. Wiley, New York.

Milinski, M., 1987. Tit for tat in sticklebacks and the evolution of cooperation. Nature 325, 433–435.

Milinski, M., Wedekind, C., 1998. Working memory constrains human cooperation in the prisoner's dilemma. Proc. Natl Acad. Sci. USA 95, 13755–13758.

Molander, P., 1985. The optimal level of generosity in a selfish, uncertain environment. J. Conflict Resolut. 29, 611–618.

Neill, D.B., 2001. Optimality under noise: higher memory strategies for the alternating prisoner's dilemma. J. Theor. Biol. 211, 159–180.

Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the prisoner's dilemma. Acta Appl. Math. 20, 247–265.

Nowak, M.A., 2006. Evolutionary Dynamics. Harvard University Press, Cambridge, MA.

Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature 359, 826–829.

Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogeneous populations. Nature 355, 250–253.

Nowak, M.A., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit for tat in prisoner's dilemma. Nature 364, 56–58.

Nowak, M.A., Sigmund, K., 1994. The alternating prisoner's dilemma. J. Theor. Biol. 168, 219–226.

Nowak, M.A., Sigmund, K., 1998. Evolution of indirect reciprocity by image scoring. Nature 393, 573–577.

Nowak, M.A., Sigmund, K., 2005. Evolution of indirect reciprocity. Nature 437, 1291–1298.

Nowak, M.A., Sigmund, K., El-Sedy, E., 1995. Automata, repeated games and noise. J. Math. Biol. 33, 703–722.

Nowak, M.A., Sasaki, A., Taylor, C., Fudenberg, D., 2004. Emergence of cooperation and evolutionary stability in finite populations. Nature 428, 646–650.

Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A., 2006. A simple rule for the evolution of cooperation on graphs. Nature 441, 502–505.

Posch, M., Pichler, A., Sigmund, K., 1999. The efficiency of adapting aspiration levels. Proc. R. Soc. London B 266, 1427–1435.

Rapoport, A., Chammah, A.M., 1965. Prisoner's Dilemma. University of Michigan Press, Ann Arbor.

Roberts, G., Sherratt, T.N., 1998. Development of cooperative relationships through increasing investment. Nature 394, 175–179.

Selten, R., Hammerstein, P., 1984. Gaps in Harley's argument on evolutionarily stable learning rules and in the logic of tit for tat. Behav. Brain Sci. 7, 115–116.

Skyrms, B., 2004. The Stag Hunt and the Evolution of Social Structure. Cambridge University Press, Cambridge, UK.

Traulsen, A., Claussen, J.C., Hauert, C., 2005. Coevolutionary dynamics: from finite to infinite populations. Phys. Rev. Lett. 95, 238701.

Traulsen, A., Pacheco, J.M., Imhof, L.A., 2006. Stochasticity and evolutionary stability. Phys. Rev. E 74, 021905.

Traulsen, A., Nowak, M.A., Pacheco, J.M., 2007. Stochastic payoff evaluation increases the temperature of selection. J. Theor. Biol. 244, 349–356.

Trivers, R.L., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35–57.

Wahl, L.M., Nowak, M.A., 1999a. The continuous prisoner's dilemma: I. Linear reactive strategies. J. Theor. Biol. 200, 307–321.

Wahl, L.M., Nowak, M.A., 1999b. The continuous prisoner's dilemma: II. Linear reactive strategies with noise. J. Theor. Biol. 200, 323–338.

Wakano, J.Y., Yamamura, N., 2001. A simple learning strategy that realizes robust cooperation better than Pavlov in iterated prisoners' dilemma. J. Ethol. 19, 1–8.

Wilkinson, G.S., 1984. Reciprocal food sharing in the vampire bat. Nature 308, 181–184.

Wilson, E.O., 1975. Sociobiology. Harvard University Press, Cambridge, MA.

Wynne-Edwards, V.C., 1962. Animal Dispersion in Relation to Social Behavior. Oliver & Boyd, Edinburgh.

Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57–84.