Forgiver Triumphs in Alternating Prisoner’s Dilemma

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation
Zagorsky, Benjamin M., Johannes G. Reiter, Krishnendu Chatterjee, and Martin A. Nowak. 2013. “Forgiver Triumphs in Alternating Prisoner’s Dilemma.” PLoS ONE 8 (12): e80814. doi:10.1371/journal.pone.0080814. http://dx.doi.org/10.1371/journal.pone.0080814.

Published Version
doi:10.1371/journal.pone.0080814

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:11879285

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Forgiver Triumphs in Alternating Prisoner’s Dilemma

Benjamin M. Zagorsky¹*, Johannes G. Reiter²*, Krishnendu Chatterjee², Martin A. Nowak¹,³*

¹ Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts, United States of America, ² IST Austria (Institute of Science and Technology Austria), Klosterneuburg, Lower Austria, Austria, ³ Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, United States of America

Abstract

Cooperative behavior, where one individual incurs a cost to help another, is a widespread phenomenon. Here we study direct reciprocity in the context of the alternating Prisoner’s Dilemma. We consider all strategies that can be implemented by one and two-state automata. We calculate the payoff matrix of all pairwise encounters in the presence of noise. We explore deterministic selection dynamics with and without mutation. Using different error rates and payoff values, we observe convergence to a small number of distinct equilibria. Two of them are uncooperative strict Nash equilibria representing always-defect (ALLD) and Grim. The third equilibrium is mixed and represents a cooperative alliance of several strategies, dominated by a strategy which we call Forgiver. Forgiver cooperates whenever the opponent has cooperated; it defects once when the opponent has defected, but subsequently Forgiver attempts to re-establish cooperation even if the opponent has defected again. Forgiver is not an evolutionarily stable strategy, but the alliance, which it rules, is asymptotically stable. For a wide range of parameter values the most commonly observed outcome is convergence to the mixed equilibrium, dominated by Forgiver. Our results show that although forgiving might incur a short-term loss it can lead to a long-term gain. Forgiveness facilitates stable cooperation in the presence of exploitation and noise.

Introduction

A cooperative dilemma arises when two cooperators receive a higher payoff than two defectors and yet there is an incentive to defect [1,2]. The Prisoner’s Dilemma [3–9] is the strongest form of a cooperative dilemma, where cooperation requires a mechanism for its evolution [10]. A mechanism is an interaction structure that specifies how individuals interact to receive payoff and how they compete for reproduction. Direct reciprocity is a mechanism for the evolution of cooperation. Direct reciprocity means there are repeated encounters between the same two individuals [11–37]. The decision whether or not to cooperate depends on previous interactions between the two individuals. Thus, a strategy for the repeated Prisoner’s Dilemma (or repeated games) is a mapping from any history of the game into what to do next. The standard theory assumes that both players decide simultaneously what to do for the next round. But another possibility is that the players take turns when making their moves [38–40]. This implementation can lead to a strictly alternating game, where the players always choose their moves in turns, or to a stochastically alternating game, where in each round the player to move is chosen at random (next is selected probabilistically). Here we investigate the strictly alternating game.

We consider the following scenario. In each round a player can pay a cost, c, for the other player to receive a benefit, b, where \( b > c > 0 \). If both players cooperate in two consecutive moves, each one gets a payoff, \( b - c \), which is greater than the zero payoff they would receive for mutual defection. But if one player defects, while the other cooperates, then the defector gets payoff, \( b \), while the cooperator gets the lowest payoff, \( -c \). Therefore, over two consecutive moves the payoff structure is the same as in a Prisoner’s Dilemma: \( b > b - c > 0 > -c \). Thus, this game is called “alternating Prisoner’s Dilemma” [29,39].

We study the strictly alternating Prisoner’s Dilemma in the presence of noise. In each round, a player makes a mistake with probability \( \varepsilon \) leading to the opposite move. We consider all strategies that can be implemented by deterministic finite-state automata [41] with one or two states. These automata define how a player behaves in response to the last move of the other player. Thus we consider a limited strategy set with short-term memory. Finite-state automata have been used extensively to study repeated games [42–45] including the simultaneous Prisoner’s Dilemma. In our case, each state of the automaton is labeled by \( C \) or \( D \). In state \( C \) the player will cooperate in the next move; in state \( D \) the player will defect. Each strategy starts in one of those two states. Each state has two outgoing transitions (either to the same or to the other state); one transition specifies what happens if the opponent has cooperated (labeled with \( c \)) and one if the opponent has defected (labeled with \( d \)). There are 26 automata encoding unique strategies (Fig. 1). These strategies include ALLC, ALLD, Grim, tit-for-tat (TFT), and win-stay lose-shift (WSLS).

ALLC (\( S_{26} \)) and ALLD (\( S_1 \)) are unconditional strategies (see Fig. 1 and Supporting File S1 for strategy names and their indexing). ALLC always cooperates while ALLD always defects.
Figure 1. Deterministic strategies in the Prisoner’s Dilemma. Each automaton defines a different strategy for how a player behaves during the game. If a player is in state C, she will cooperate in the next move; if she is in state D, then she will defect. The outgoing transitions of a state define how the state of an automaton will change in response to cooperating (label c) or defecting (label d) of the opponent. The left state with the

Suspicious dynamic strategies

- **S2**: Suspicious Paradoxic (S3)
  - Transition: C → D, D → C
- **S6**: Suspicious Alternator (S10)
  - Transition: C → D, D → C
- **S7**: Suspicious WSLS (S4)
  - Transition: C → D, D → C
- **S11**: Suspicious TFT (S8)
  - Transition: C → D, D → C
- **S12**: Suspicious Forgive (S12)
  - Transition: C → D, D → C

Sink-state C strategies

- ALLC (S26)
  - Transition: C → C, D → C
- Paradoxic Grateful (S5)
  - Transition: C → C, D → C
- Grateful (S9)
  - Transition: C → C, D → C
- Suspicious ALLC (S13)
  - Transition: C → C, D → C

Hopeful dynamic strategies

- Forgiver (S14)
  - Transition: C → D, D → C
- TFT (S15)
  - Transition: C → D, D → C
- WSLS (S16)
  - Transition: C → D, D → C
- Alternator (S22)
  - Transition: C → D, D → C
- S18
  - Transition: C → D, D → C
- S19
  - Transition: C → D, D → C
- Paradoxic (S20)
  - Transition: C → D, D → C
- S23
  - Transition: C → D, D → C
- S24
  - Transition: C → D, D → C
- S25
  - Transition: C → D, D → C
- ALLD (S1)
  - Transition: C → D, D → C

Sink-state D strategies

- Grim (S17)
  - Transition: C → D, D → C
- Paradoxic Grim (S21)
  - Transition: C → D, D → C
- Hopeful ALLD (S25)
  - Transition: C → D, D → C
- ALLD (S1)
  - Transition: C → D, D → C
Both strategies are implemented by a one-state automaton (Fig. 1). The strategy Grim starts and stays in state C as long as the opponent cooperates. If the opponent defects, Grim permanently moves to state D with no possibility to return. TFT (S15) starts in state C and subsequently does whatever the opponent did in the last round [5]. This simple strategy is very successful in an error-free environment as it promotes cooperative behavior but also avoids exploitation by defectors. However, in a noisy environment TFT achieves a very low payoff against itself since it can only avoid exploitation by defectors. However, in a noisy environment as it promotes cooperative behavior but also avoids exploitation by defectors. However, in a noisy environment if the opponent defect, it moves to the defection state with no possibility of recovering from an accidental defection within three rounds. Against a copy of itself, Forgiver performs very well as it can recover from an accidental defection within three rounds. Against defection-heavy strategies like Grim and ALLD, Forgiver gets exploited in each second round. Both TFT and WSLS are not error correcting as they are unable to recover back to cooperation after an unintentional mistake. Only another mistake can enable them to return to cooperative behavior. When Grim plays against itself and a single defection occurs, it moves to the defection state with no possibility of returning to cooperation.

A stochastic variant of Forgiver is already described in [39]. In this study, strategies are defined by a quadruple \((p_1, p_2, p_3, p_4)\) where \(p_i\) denotes the probability to cooperate after each of the four outcomes CC, CD, DC, and DD. This stochastic strategy set is studied in the setting of the infinitely-repeated alternating game. The initial move is irrelevant. In [39] a strategy close to Forgiver (S16) is victorious in computer simulations of the strictly alternating Prisoner’s Dilemma. For further discussions see also pp. 78–80 in [29]; there the stochastic variant of Forgiver is called ‘Firm but Fair’.

**Results**

We calculate the payoff for all pairwise encounters in games of \(L\) moves of both strategies, thereby obtaining a 26 × 26 payoff matrix. We average over which strategy goes first. Without loss of generality we set \(c = 1\). At first we study the case \(b = 2\) with error rate \(\epsilon = 0.05\) and an average game length of \(L = 100\). Table 1 shows a part of the calculated payoff matrix for six relevant strategies. We find that ALLD (S1) and Grim (S17) are the only strict Nash equilibria among the 26 pure strategies. ALLC (S26) vs ALLC receives a high payoff, but so does Forgiver vs Forgiver. The payoffs of WSLS vs WSLS and TFT vs TFT are low, because
neither strategy is error correcting (Fig. 2). Interestingly TFT vs WSLS yields good payoff for both strategies, because their interaction is error correcting.

In the following, we study evolutionary game dynamics [48–50] with the replicator equation. The frequency of strategy \( S_i \) is denoted by \( x_i \). At any one time we have \( \sum_{i=1}^{n} x_i = 1 \), where \( n = 26 \) is the number of strategies. The frequency \( x_i \) changes according to the relative payoff of strategy \( S_i \). We evaluate evolutionary trajectories for many different initial frequencies. The trajectories start from \( 10^4 \) uniformly distributed random points in the 26-simplex.

Typically, we do not find convergence to one of the strict Nash equilibria (Fig. 3 b). In only 5% of the cases the trajectories converge to the pure ALLD equilibrium and in 18% of the cases the trajectories converge to the pure Grim equilibrium. However, in 77% of the cases we observe convergence to a mixed equilibrium of several strategies, dominated by Forgiver with a population share of 82.6% (Fig. 3 b). The other six strategies present in this cooperative simplex.

Table 1. Payoff matrix for the most relevant strategies.

|       | ALLD | Forgiver | TFT | WSLS | Grim | ALLC |
|-------|------|---------|-----|------|------|------|
| ALLD  | 10.0 | 148.4   | 24.8| 144.9| 11.5 | 280.0|
| Forgiver | -36.1| 174.8   | 163.5| 166.9| -12.7| 194.3|
| TFT   | 5.1 | 178.1   | 104.5| 176.7| 24.5 | 194.5|
| WSLS  | -35.0| 169.1   | 162.3| 106.5| -12.0| 230.9|
| Grim  | 9.5 | 152.0   | 40.5| 148.8| 28.1 | 262.9|
| ALLC  | -80.0| 177.2   | 176.6| 67.3 | -28.8| 190.0|

Excerpt of the payoff matrix with the most relevant strategies when the benefit value \( b = 3 \) (\( c = 1 \)), the error rate \( \epsilon = 5\% \), and the number of rounds in each game \( L = 100 \). There are two pure Nash equilibria in the full payoff matrix: ALLD (\( S_1 \)) and Grim (\( S_{17} \)), both denoted in bold.

doi:10.1371/journal.pone.0080814.t001

For the calculation of average payoff per round for the most relevant strategy pairs (Table 3; for the calculations see File S1: Section 2 and Fig. S8–S10). From these results we obtain that ALLD (or ssD strategies) cannot invade Forgiver if

\[
\frac{b}{c} > 2 + \epsilon - \epsilon^2 - \frac{1}{1 - 2\epsilon}
\]

This result holds for any error rate, \( \epsilon \), between 0 and 1/2 (Fig. 4 d).

Discussion

Our results imply an indubitable strength of the strategy Forgiver in the alternating Prisoner’s Dilemma in the presence of noise. For a wide range of parameter values, Forgiver is the dominating strategy of the cooperative equilibrium, having a population share of more than half in all investigated scenarios.

Essential for the success of a cooperative strategy in the presence of noise is how fast it can recover back to cooperation after a mistake, but at the same time, also avoid excessive exploitation by defectors. The conditional cooperation element is crucial for the triumph of Forgiver. Even though, also TFT and WSLS contain this element, which allows them to cooperate against cooperative strategies without getting excessively exploited by defectors, these strategies are not as successful as Forgiver, because of their inability to correct errors. Grim also possesses this conditional cooperation element. However, noise on the part of Grim’s opponent will inevitably cause Grim to switch to always-defect. It is Grim’s ability to conditionally cooperate for the first handful of mistakes that provides a competitive advantage over pure ALLD such that the strict Nash equilibrium ALLD can only rarely arise.

The other strategies appearing in the Forgiver equilibrium for the cases of \( b = 3, b = 4 \), and \( b = 5 \) are Paradoxical Grateful (\( S_5 \)), Grateful (\( S_6 \)), Suspicious ALLC (\( S_{13} \)) and ALLC (\( S_{26} \)). All of them are ssC strategies, that, in the presence of noise, behave like ALLC against Grim can become higher than the payoff of Grim against itself. In other words, Grim can be invaded by ALLC. Hence, instead of the pure Grim equilibrium we observe a mixed equilibrium between Grim and ALLC (Fig. S4b–d in File S1).

We check the robustness of the observed equilibria by incorporating mutation to the replicator equation. We find that both the ALLD and the rare Suspicious Forgiver equilibrium are unstable. In the presence of mutation the evolutionary trajectories lead away from ALLD to Grim and from Suspicious Forgiver to Forgiver (see Fig. S6 and S7 in File S1). The Grim equilibrium and the Forgiver equilibrium remain stable. We note that this asymptotic stability is also due to the restricted strategy space. In [51] it has been shown that in the simultaneous Prisoner’s Dilemma with an unrestricted strategy space, no strategy is robust against indirect invasions and hence, no evolutionarily stable strategy can exist.

Essential for the stability in our model is that Forgiver can resist invasion by ssD strategies (\( S_1, S_7, S_{11}, S_{27} \)), because Forgiver does better against itself than the ssD strategies do against Forgiver (Table 1). However, Forgiver can be invaded by ssC strategies and TFT. But, since TFT performs poorly against itself and ssC strategies are exploited by WSLS (Table 1), all these strategies can coexist in the Forgiver equilibrium. Stable alliances of cooperative strategies have also been found in the context of the Public Goods Game [52] and indirect reciprocity [53]. More detailed results and equilibrium analysis for a wide range of parameter values for \( \epsilon \) and \( L \) are provided in File S1 (Tables S1–S14 and Figures S2–S7).

In the limit of infinitely many rounds per game, we can use an infinitely repeated game, and we derive analytical results for the average payoff per round for the most relevant strategy pairs (Table 3; for the calculations see File S1: Section 2 and Fig. S8–S10). From these results we obtain that ALLD (or ssD strategies) cannot invade Forgiver if

\[
\frac{b}{c} > 2 + \epsilon - \epsilon^2 - \frac{1}{1 - 2\epsilon}
\]

This result holds for any error rate, \( \epsilon \), between 0 and 1/2 (Fig. 4 d).
after the first few moves. The strategy ALLC does very well in combination with Forgiver. Nevertheless, ALLC itself appears rarely. Perhaps because of Paradox Grateful, which defects against ALLC for many moves in the beginning, whereas Suspicious ALLC puts Paradox Grateful into its cooperating state immediately. One might ask why these ssC strategies do not occupy a larger population share in the cooperative equilibrium. The reason is the presence of exploitative strategies like WSLS which itself is a weak strategy in this domain. If only Forgiver was present, WSLS would be quickly driven to extinction; WSLS does worse against itself and Forgiver than Forgiver does against WSLS and itself (see Table 1). But WSLS remains in the Forgiver equilibrium because it exploits the ssC strategies. Interestingly, higher error rates can decrease the probability to converge to the cooperative equilibrium dramatically and hence prevent the evolution of any cooperative behavior (Fig. 4a).

Grim and Forgiver are similar strategies, the difference being, in the face of a defection, Forgiver quickly returns to cooperation whereas Grim never returns. An interesting interpretation of the relationship is that Grim never forgives while Forgiver always does. Thus, the clash between Grim and Forgiver is actually a test of the viability of forgiveness under various conditions. On the one hand, the presence of noise makes forgiveness powerful and essential. On the other hand, if cooperation is not valuable enough, forgiveness can be exploited. Moreover, even when cooperation is valuable, but the population is ruled by exploiters, forgiveness is not a successful strategy. Given the right conditions, forgiveness makes cooperation possible in the face of both exploitation and noise.

These results demonstrate a game-theoretic foundation for forgiveness as a means of promoting cooperation. If cooperation is valuable enough, it can be worth forgiving others for past wrongs in order to gain future benefit. Forgiving incurs a short-term loss but ensures a greater long-term gain. Given all the (intentional or unintentional) misbehavior in the real world, forgiveness is essential for maintaining healthy, cooperative relationships.

Methods

Strategy space

We consider deterministic finite automata [41] (DFA) with one and two states. There are two one-state automata which encode
the strategies always-defect (ALLD) and always-cooperate (ALLC).
In total, there are 32 two-state automata encoding strategies in our game: two possible arrangements of states (CD, DC) and 16 possible arrangements of transitions per arrangement of states. For 8 of these 32 automata, the second state is not reachable, making them indistinguishable from a one-state automata. Since we already added the one-state automata to our strategy space, these 8 can be ignored. The remaining 24 two-state automata encode

Figure 4. Robustness of results across various benefit values and error rates. a | Convergence probability to the Forgiver equilibrium of a uniform-random point in the 26-simplex. Note that for higher error rates (increasing noise-level), the probability to converge to the cooperative equilibrium is much lower. b | Population share of Forgiver (S₁₄) in the Forgiver equilibrium. Observe the relationship between the higher error rates and the lower population share of Forgiver. c | Population share of sink-state C strategies (S₅, S₉, S₁₃, S₂₆) in the Forgiver equilibrium. Higher error rates lead to higher proportions of unconditional cooperators. d | In the infinitely repeated game, for all value pairs of b/c and ε in the blue shaded area, ALLD cannot invade Forgiver since the average payoff of Forgiver playing against itself is higher than the average payoff of ALLD against Forgiver (see Inequality (1)).

doi:10.1371/journal.pone.0080814.g004

Table 2. Equilibrium frequencies.

|     | ALLD | Grim | Forgiver | sForgiver |
|-----|------|------|----------|-----------|
| b = 2 | 15%  | 52%  | 33%      | <1%       |
| b = 3 | 5%   | 18%  | 77%      | 0%        |
| b = 4 | 2%   | 7%   | 90%      | 1%        |
| b = 5 | 1%   | 3%   | 93%      | 3%        |

Proportions in which the four equilibria were reached from 10⁹ uniformly distributed random points in the 26-simplex. In the case of b = 2, the mixed Forgiver equilibrium has a different composition than in the cases of b = 3, b = 4, b = 5. Parameter values: costs c₁ = 1, error rate ε = 0.05, number of rounds per game L = 100.

doi:10.1371/journal.pone.0080814.t002

Table 3. Analytical results in the infinitely alternating Prisoner’s Dilemma.

|     | ALLD | Forgiver | ALLC |
|-----|------|----------|------|
| ALLD | c(b - c) | \(\frac{1-c^2}{2} - c\) | \(b - c(b + c)\) |
| Forgiver | \(c - \frac{1-c^2}{2} - c\) | \(1 + c^2(1-\epsilon)\) | \(b(1-\epsilon) - c(1-c)\) |
| ALLC | \(c - \frac{1-c^2}{2} - c\) | \(1 + c^2(1-\epsilon)\) | \(b(1-\epsilon) - c(1-c)\) |

Analytical results of the average payoff per round in the infinitely alternating Prisoner’s Dilemma for ALLD (S₁), Forgiver (S₁₄), and ALLC (S₂₆) playing against each other. Derivations are provided in File S1 (section 2).

doi:10.1371/journal.pone.0080814.t003
distinct strategies in our game. Hence, in total we have 26 deterministic strategies in the alternating Prisoner’s Dilemma (Fig. 1).

**Generation of the payoff matrix**

In each round of the game a player can either cooperate or defect. Cooperation means paying a cost, \( c \), for the other player to receive a benefit, \( b \). Defection means paying no cost and distributing no benefit. If \( b > c > 0 \) and we sum over two consecutive moves (equivalent to one round), the game is a Prisoner’s Dilemma since the following inequality is satisfied: \( b > b - c > 0 > -c \). In other words, in a single round it is best to defect, but cooperation might be fruitful when playing over multiple rounds. Furthermore also \( 2(b - c) > b - c > 0 \) holds, and hence mutual cooperation results in a higher payoff than alternating between cooperation and defection. The second inequality ensures that sustained cooperation results in a higher payoff than alternation between cooperation and defection.

For each set of parameters (number of rounds \( L \), error rate \( \epsilon \), benefit value \( b \), and costs \( c \)), we generate a \( 26 \times 26 \) payoff matrix \( A \) where each of the 26 distinct strategies is paired with each other. The entry \( a_{ij} \) in the payoff matrix \( A \) gives the payoff of strategy \( S_i \) playing against strategy \( S_j \). Based on the average of which strategy (player) goes first, we define the initial state distribution of both players as a row vector \( Q_{S_i,S_j} \). Since the players do not observe when they have made a mistake (i.e., the faulty player does not move to the corresponding state of the erroneous action which he has accidentally played), the state space consists of sixteen states namely \( CC, CD, DC, DD, D'C, D'D, C'C, C'D, CD', \ldots, C'C' \). The star after a state indicates that the player accidentally played the opposite move as intended by her current state.

Each game consists of \( L \) moves of both player. In each move, a player makes a mistake with probability \( \epsilon \) and thus implements the opposite move of what is specified by her strategy (automaton). We denote \( 1 - \epsilon \) by \( \bar{\epsilon} \). Although, the players do not observe their mistakes, the payoffs depend on the actual moves. This setting relates to imperceptive implementation errors [16,18,29,45] (see section 3 in File S1 for a discussion on error types). The payoffs corresponding to their moves in the different states are given by the column vector \( U \).

Next, we define a \( 16 \times 16 \) transition matrix \( M_{S_i,S_j} \) for each pair of strategies \( S_i, S_j \). The entries of the transition matrix are given by the probabilities to move from each state of the sixteen states (defined above) to the next:

\[
M_{S_i,S_j} = \begin{pmatrix}
p_1 p_2 x_1 & (1-p_1) p_2 x_2 & (1-p_2) p_2 x_3 & (1-p_2) (1-p_2) x_4 & \ldots & (1-p_1)(1-p_2)p_2 x_4 \\
p_2 p_2 x_1 & \vdots & \vdots & \vdots & \ldots & \vdots \\
p_2 x_2 & \vdots & \vdots & \vdots & \ldots & \vdots \\
p_2 x_3 & \vdots & \vdots & \vdots & \ldots & \vdots \\
p_2 x_4 & \vdots & \vdots & \vdots & \ldots & \vdots \\
\end{pmatrix}
\]

where the quadruple [39] \((p_1, p_2, p_3, p_4)\) defines the probabilities of strategy \( S_j \) to cooperate in the observed states \( CC, CD, DC, \) and \( DD \) (errors remain undetected by the players). Respectively, the quadruple \( (p'_1, p'_2, p'_3, p'_4) \) encodes the strategy \( S_j \). For example, \((1-p_3)p_2x_3 \) is the probability to move from state \( DD \) to state \( C'C \). A deterministic strategy is represented as a quadruple where each \( p_i \in \{0,1\} \).

Using the initial state distribution \( Q_{S_i,S_j} \), the transition matrix \( M_{S_i,S_j} \), and the payoff vector \( U \), we calculate the payoff \( a_{ij} \) of strategy \( S_i \) playing against strategy \( S_j \) via a Markov Chain:

\[
a_{ij} = Q_{S_i,S_j} \sum_{k=0}^{L-1} M_{k,S_i,S_j} U .
\]

Applying equation (3) to each pair of strategies, we obtain the entire payoff matrix \( A \) for a given set of parameter values. Although we use deterministic strategies, the presence of noise implies that the game that unfolds between any two strategies is described by a stochastic process. Payoff matrices for benefit values of \( b = 2, b = 3, b = 4, \) and \( b = 5 \), for error rates of \( \epsilon = 0.01, \epsilon = 0.05, \) and \( \epsilon = 0.1, \) and for game length of \( L = 10, L = 100, \) and \( L = 1000 \) are provided in File S1 (Tables S1–S8).

**Evolution of strategies.** The strategy space spans a 26-simplex which we explore via the replicator equation [48–50] with and without mutations. The frequency of strategy \( S_i \) is given by \( x_i \). At any time \( \sum_{i=1}^{n} x_i = 1 \) holds where \( n = 26 \) is the number of strategies. The average payoff (fitness) for strategy \( S_i \) is given by

\[
f_i = \sum_{j=1}^{n} a_{ij} x_j .
\]

The frequency of strategy \( S_i \) changes according to the differential equation

\[
\dot{x}_i = x_i (f_i - \bar{f}) + u \left( \frac{1}{n} - x_i \right)
\]

where the average population payoff is \( \bar{f} = \sum_{i=1}^{n} f_i x_i \) and \( u \) is the mutation rate. Mutations to each strategy are equally likely; for non-uniform mutation structures see [54]. Using the differential equation (5), defined on the \( n \)-simplex [here, \( n = 26 \)], we study the evolutionary dynamics in the alternating Prisoner’s Dilemma for many different initial conditions (i.e., random initial frequencies of the strategies). We generate a uniform-random point in the \( n \)-simplex by taking the negative logarithm of \( n \) random numbers in \((0,1)\), then normalizing these numbers such that they sum to 1, and using the normalized values as the initial frequencies of the \( n \) strategies [55].

**Computer simulations.** Our computer simulations are implemented in Python and split into three programs. The first program generates the \( 26 \times 26 \) payoff matrix for each set of parameters. The second program simulates the deterministic selection dynamics starting from uniform-random points in the 26-simplex. The third program performs statistical analysis on the results of the second program. The code is available at http://pub.ist.ac.at/~jevreter upon request [56].

**Supporting Information**

File S1 Detailed description of the model and the strategies; Simulation results and equilibrium analysis for a wide range of parameter values; Calculations for the infinitely-repeated game; Implementation of errors; Includes Tables S1–S14 and Figures S1–S10. (PDF)

**Acknowledgments**

We thank Sasha Rubin and Sebastian Novak for helpful comments and suggestions.
Author Contributions
Conceived and designed the experiments: BMZ JGR MAN. Performed the experiments: BMZ JGR MAN. Analyzed the data: BMZ JGR MAN. Wrote the paper: BMZ JGR MAN. Provided input to the manuscript: KC.

References
1. Hauert C, Michor F, Nowak MA, Doebeli M (2006) Synergy and discounting of cooperation in social dilemmas. Journal of Theoretical Biology 239: 195–202.
2. Nowak MA (2012) Evolving cooperation. Journal of Theoretical Biology 299: 1–5.
3. Rapoport A, Chammah AM (1965) Prisoner’s dilemma. University of Michigan Press.
4. Trivers RL (1971) The evolution of reciprocal altruism. Quarterly review of biology 46: 35–57.
5. Axelrod R (1984) The Evolution of Cooperation. New York: Basic Books.
6. May RM (1987) More evolution of cooperation. Nature 327: 15–17.
7. Milinski M (1987) Tit for tat in sticklebacks and the evolution of cooperation. Nature 325: 433–435.
8. Sigmund K (1993) Games of life: explorations in ecology, evolution and behavior. Oxford University Press, Oxford.
9. Clutton-Brock T (2009) Cooperation between non-kin in animal societies. Nature 462: 51–57.
10. Nowak MA (2006) Five rules for the evolution of cooperation. Science 314: 1560–1563.
11. Aumann RJ (1981) Survey of repeated games. In: Essays in game theory and mathematical economics in honor of Oskar Morgenstern, Bibliographisches Institut Mannheim. pp. 11–42.
12. Fudenberg D, Maskin E (1986) The folk theorem in repeated games with discounting or with incomplete information. Econometrica: Journal of the Econometric Society 54: 533–554.
13. Abreu D (1988) On the theory of infinitely repeated games with discounting. Econometrica 56: 383–396.
14. Kraines D, Kraines V (1982) Pavlov and the prisoner’s dilemma. Theory and Decision 26: 47–79.
15. Nowak MA, Sigmund K (1989) Oscillations in the evolution of reciprocity. Journal of Theoretical Biology 137: 21–26.
16. Fudenberg D, Maskin E (1990) Evolution and cooperation in noisy repeated games. The American Economic Review 80: 274–279.
17. Nowak MA, Sigmund K (1990) The evolution of stochastic strategies in the prisoner’s dilemma. Acta Applicandae Mathematicae 20: 247–265.
18. Wu J, Axelrod R (1995) How to cope with noise in the iterated prisoner’s dilemma game. Journal of Economic Theory 68: 234–248.
19. Kraines D, Kraines V (1995) Evolution of learning among pavlov strategies in a competitive environment with noise. Journal of Conicet Resolution 39: 439–466.
20. Cressman R (1996) Evolutionary stability in the finitely repeated prisoner’s dilemma. Journal of Economic Theory 68: 234–248.
21. Boeroijst MC, Nowak MA, Sigmund K (1997) Equal pay for all prisoners. American Mathematical Monthly 104: 303–305.
22. Milinski M, Wedekin D (1998) Working memory constrains human cooperation in the prisoners dilemma. Proc Natl Acad Sci USA 95: 13575–13578.
23. Kraines DP, Kraines VY (2000) Natural selection of memory-one strategies for the iterated prisoners dilemma. Journal of Theoretical Biology 203: 333–355.
24. Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. Nature 428: 646–650.
25. Dal Bo P (2005) Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. The American Economic Review 95: 1591–1604.
26. Doebeli M, Hauert C (2005) Models of cooperation based on the prisoner’s dilemma and the snowdrift game. Ecology Letters 8: 748–766.
27. Kendall G, Yao X, Chong SY (2007) The Iterated Prisoners’ Dilemma: 20 Years on. World Scientific Publishing Co., Inc.
28. Chen X, Fu F, Wang L (2008) Interaction stochasticity supports cooperation in spatial prisoners dilemma. Physical Review E 78: 051120.
29. Sigmund K (2009) The calculus of selfishness. Princeton University Press.
30. Perc M, Szolnoki A (2010) Coevolutionary games—a mini review. Biosystems 99: 109–125.
31. Imhof LA, Nowak MA (2010) Stochastic evolutionary dynamics of direct reciprocity. Proceedings of the Royal Society 277: 463–466.
32. Fudenberg D, Rand DG, Dreher A (2012) Slow to anger and fast to forgive: cooperation in an uncertain world. American Economic Review 102: 720–749.
33. van Veelen M, Garcia J, Rand DG, Nowak MA (2012) Direct reciprocity in structured populations. Proc Natl Acad Sci USA 109: 9929–9934.
34. Press WH, Dyson FJ (2012) Iterated prisoners dilemma contains strategies that dominate any evolutionary opponent. Proc Natl Acad Sci USA 109: 10409–10413.
35. Wang Z, Szolnoki A, Perc M (2013) Interdependent network reciprocity in evolutionary games. Scientific Reports 3.
36. Hilbe C, Nowak MA, Sigmund K (2013) The evolution of extortion in iterated prisoner’s dilemma games. Proc Natl Acad Sci USA 110: 6913–6918.
37. Stewart AJ, Plotkin JB (2013) From extortion to generosity, evolution in the iterated prisoners dilemma. Proc Natl Acad Sci USA 110: 15348–15353.
38. Wedekind C, Milinski M (1996) Human cooperation in the simultaneous and the alternating prisoner’s dilemma: Pavlov versus generous tit-for-tat. Proc Natl Acad Sci USA 93: 2686–2689.
39. Nowak MA, Sigmund K (1994) The alternating prisoners dilemma. Journal of Theoretical Biology 166: 219–226.
40. Frean MR (1994) The prisoner’s dilemma without synchrony. Proceedings of the Royal Society of London 257: 75–79.
41. Hauert C, Matsumo R, Sigmund K (2006) Introduction to Automata Theory, Languages, and Computation. Prentice Hall, 3rd edition.
42. Rubinstein A (1986) Finite automata play the repeated prisoner’s dilemma. Journal of Economic Theory 39: 83–96.
43. Miller JH (1996) The coevolution of automata in the repeated prisoner’s dilemma. Journal of Economic Behavior & Organization 29: 87–112.
44. Bimbo KG, Samuelson L (1992) Evolutionary stability in repeated games played by finite automata. Journal of Economic Theory 57: 378–395.
45. Nowak MA, Sigmund K, El-Sheh E (1995) Automata, repeated games and noise. Journal of Mathematical Biology 33: 703–722.
46. Nowak MA, Sigmund K (1992) Tit for tat in heterogeneous populations. Nature 355: 250–253.
47. Nowak MA, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game. Nature 364: 36–38.
48. Hofbauer J, Sigmund K (1998) The Theory of Evolution and Dynamical Systems. Mathematical Aspects of Selection. Cambridge University Press.
49. Hofbauer J, Sigmund K (1998) Evolutionary Game Theory. The MIT press.
50. van Veelen M, Garcia J (2010) In and out of equilibrium: Evolution of strategies in repeated games with discounting. T1 discussion paper 10-037 / 1.
51. Szolnoki A, Perc M (2010) Reward and cooperation in the spatial public goods game. EPL 92: 38003.
52. Brantl H, Sigmund K (2005) Indirect reciprocity, image scoring, and moral hazard. Proc Natl Acad Sci USA 102: 2666–2670.
53. Garcia J, Traulsen A (2012) The structure of mutations and the evolution of reciprocity. Proceedings of the Royal Society 277: 1591–1604.
54. URL http://pub.ist.ac.at/~jreiter. Personal website of JG (Accessed Oct. 2013).