ORIGINAL RESEARCH

Asymptotic behavior of strategies in the repeated prisoner's dilemma game in the presence of errors

Christos A. Ioannou *

Department of Economics, University of Southampton, United Kingdom

Received: July 31, 2014 **Accepted:** September 17, 2014 **Online Published:** September 28, 2014
DOI: 10.5430/air.v3n4p28 **URL:** http://dx.doi.org/10.5430/air.v3n4p28

Abstract

We examine the asymptotic behavior of a finite, but error-prone population, whose agents can choose one of ALLD (always defect), ALLC (always cooperate), or Pavlov (repeats the previous action if the opponent cooperated and changes action otherwise) to play the repeated Prisoner's Dilemma. A novelty of the study is that it allows for three types of errors that affect agents' strategies in distinct ways: (a) implementation errors, (b) perception errors of one's own action, and (c) perception errors of the opponent's action. We also derive numerical results based on the payoff matrix used in the tournaments of Axelrod. Strategies' payoffs are monitored as the likelihood of committing errors increases from zero to one, which enables us to provide a taxonomy of best response strategies. We find that for some range of error levels, a unique best response (i.e. a dominant strategy) exists. In all other, the population composition can vary based on the proportion of each strategist's type and/or the payoffs of the matrix. Overall, our results indicate that the emergence of cooperation is considerably weak at *most* error levels.

Key Words: Prisoner's Dilemma, Bounded Rationality, Markov process, Dominant strategies

1 Introduction

Agents engage in behaviors that are constrained by the limitations of their nature and the surrounding environment. Such limitations have been treated by researchers under the rubric of "errors." Errors oftentimes result in unintended actions and/or incorrect inferences, which might lead to significant complications quite fast. For example, on 21 January, 1968, a nuclear-armed United States Air Force B-52 aircraft on a Cold War "Chrome Dome" mission, crushed near Thule Air Base in the Danish-administered territory of Greenland. Operation "Chrome Dome," initiated in 1960, was one of several United States Air Force Cold-War era airborne global alert programs in which B-52 bomber aircrafts armed with thermonuclear weapons were assigned targets in the U.S.S.R. on schedules guaranteeing that a substantial number of them were flying and fueled for their missions at any given time. The incident led to a short escalation of Cold War tensions between the Americans and Soviets. A much sharper escalation occurred on September 1, 1983, when Korean Air Lines Flight 007 was shot down by the Soviets, killing all 269 people aboard, after the aircraft strayed into prohibited Soviet airspace around the time of a planned missile test. Soviets initially denied knowledge of the incident, but later admitted the shoot-down, claiming that the aircraft was on a spy mission.

Our objective in this study is to examine the asymptotic behavior of a finite, but error-prone population of agents that play the repeated Prisoner's Dilemma game. The latter game has become the theoretical gold standard for investigating interactions; its importance stems from defying common sense reasoning and highlighting the omnipresent

* **Correspondence:** Christos A. Ioannou; Email: christos.a.ioannou@gmail.com; Address: Department of Economics, University of Southampton, United Kingdom

conflict of interests among unrelated agents. The generic Prisoner's Dilemma game is indicated in Table 1. Furthermore, we restrict significantly the number of strategies considered to ensure that the analysis of the selection dynamics is tractable. We thus concentrate on three, memory-one strategists: the *unconditional egoist*, who always defects (ALLD); the *unconditional altruist*, who always cooperates (ALLC); and the *opportunist*, who repeats the previous action if the opponent cooperated and changes action otherwise (Pavlov).

Table 1: Generic Prisoner's Dilemma Payoff Matrix

	Cooperate	Defect
Cooperate	R	S
Defect	T	P

Note. Agents either *Cooperate* (C) or *Defect* (D). If both agents choose C, then each earns the *Reward* payoff (R). If both agents choose D, then each earns the *Punishment* payoff (P). If one agent chooses D and the other agent chooses C, then the former agent earns the *Temptation* payoff (T), and the latter agent earns the *Sucker's* payoff (S). The payoffs are ordered such that $T>R>P>S$ and satisfy $R>$ T+S/2. The payoffs are those of the row agent.

A novelty of the study is that we allow for three types of errors that affect agents' strategies in distinct ways. The first channel models errors in the implementation of actions along the lines of Selten's trembling hand.[2] The other two channels model errors in the transmission of information; in particular, the second channel models errors in the transmission of one's own action, whereas the third channel models errors in the transmission of the opponent's action. For instance, the first incident in the opening paragraph is an example of an implementation error, whereas the second incident is an example of a perception error of the opponent's action. Crucially, the error types allowed affect the strategies studied in different ways. Neither ALLC nor ALLD are affected by perception errors of one's own action or perception errors of the opponent's action. Yet both strategies are affected by implementation errors. On the contrary, Pavlov is affected by all three types of errors. Even though many studies investigate the impact of errors on selected strategies, to the best of our knowledge, this is the first study to allow for all three types of errors. Other studies that investigate the impact of errors on strategies, albeit consider only implementation errors or perception errors of the opponent's action, are those of Molander,[3] Fudenberg and Maskin,[4] Nowak et al,[5] Kraines and Kraines,[6] Wu and Axelrod,[7] Wahl and Nowak,[8] Kraines and Kraines,[9] Panchanathan and Boyd,[10] Nowak and Sigmund,[11] Imhof et al.,[12] and Rand et al.[13]

We model the successive actions chosen by the agents using such strategies by a Markov process. Moreover, the presence of errors guarantees that the process is ergodic. We thus compute the invariant distributions of the realized actions for each pair of strategies and obtain the asymptotic payoff matrix across the three strategies. We also derive numerical results based on the payoff matrix used in the celebrated tournaments of Axelrod.[1] Our preference towards using the specific payoff matrix is twofold: first, Axelrod's tournaments were error-free, therefore it would be interesting to investigate the interplay of these three simple, but fundamental strategies in the presence of errors, and second, there is a vast literature that succeeded the findings in the tournaments of Axelrod, which can serve as a motivation to the discussion of this study. Strategies' payoffs are monitored as the likelihood of committing errors increases from zero to one. Traditionally, error levels have been assumed quite small. This is a plausible assumption to make in many environments, but *not* in every environment − for instance, in environments where agents lack vital resources, the likelihood of committing errors is quite high. We thus propose a systematic analysis of the entire range of error levels to obtain a taxonomy of best response strategies in the presence of errors. When a unique best response is indicated, then that best response is also a dominant strategy. Consequently, the taxonomy enables us to also determine dominant strategies.

We find that for some range of error levels, a dominant strategy exists. In the error levels where a dominant strategy does exist, ALLD is the dominant strategy at low error levels; Pavlov is the dominant strategy at intermediate levels; and ALLC, which acts as if it were an ALLD in low error levels, is the dominant strategy at high error levels. The notion of dominance is of paramount importance to the selection dynamics. More specifically, an ecological selection process will enable the dominant strategy to proliferate to the point where the entire population converges to a pure one implementing that strategy. Recall that in the ecological perspective, there is a changing distribution of the strategists' types. The less successful strategists become less common and the more successful strategists proliferate. The latter perspective differs from an evolutionary perspective, which would allow mutations to introduce new strategies into the environment. Furthermore, in the error levels where a dominant strategy does not exist, the population composition can vary based on the proportion of each strategist's type and/or the payoffs of the Prisoner's Dilemma matrix. Overall, our results indicate that the emergence of cooperation is considerably weak at *most* error levels.

The rest of the paper is organized as follows. In Section 2, we describe the analytical framework using a Markov process. In Section 3, we revisit Axelrod's tournaments. We first review some important findings highlighted in the celebrated tournaments, and then discuss subsequent seminal studies. In Section 4, we use the numerical values of Axelrod's payoff matrix to derive and display graphically the payoffs of the strategies in the head-to-head competitions for the entire range of errors. We also present examples to discuss the implications for the population dynamics and provide a taxonomy of best response strategies in the presence of errors. In Section 5, we discuss the important findings. Finally, in Section 6, we offer concluding remarks and direction for future research.

2 Markov process

We provide next the framework to study the asymptotic behavior of a finite, but error-prone population, whose agents can choose one of ALLD, ALLC, or Pavlov to play the repeated Prisoner's Dilemma game. Each period of play leads to an outcome j ($j = 1, 2, 3, 4$): (C,C), (C,D), (D,C), and (D,D). Note that the first position denotes the action taken by agent i and the second position that of agent $-i$. The transition rules are labeled by quadruples (s_1, s_2, s_3, s_4) of zeros and ones. In this context, s_j is 1 if the strategy plays $Cooperate$ and 0 if the strategy plays $Defect$, after outcome j is realized. The transition rules for ALLD, ALLC, and Pavlov are $(0, 0, 0, 0)$, $(1, 1, 1, 1)$, $(1, 0, 0, 1)$, respectively. Consider the transition rule for Pavlov. Recall that Pavlov repeats the previous action if the opponent cooperated and changes action otherwise. The transition rule signifies that the probability of cooperating is 1, if the outcome (C,C) or (D,D) is realized. Otherwise, if outcome (C,D) or (D,C) is observed, then the probability of cooperating is 0. For convenience, these rules are labeled S^{ALLD}, S^{ALLC}, and S^{Pavlov}. Suppose that strategies are subjected to three types of errors: (a) implementation errors, (b) perception errors of one's own action, and (c) perception errors of the opponent's action. Let ϵ denote the probability of committing an implementation error, δ the probability of committing a perception error of the opponent's action, and ζ the probability of committing a perception error of one's own action.

A stochastic strategy has transition rules $\mathbf{p} = (p_1, p_2, p_3, p_4)$, where p_j is any number between 0 and 1 denoting the $probability\ of\ cooperating$ after the corresponding outcome of the previous period. The space of all such rules is the four-dimensional unit cube; the corners are just the degenerate transition rules. In the context of the proposed framework, the stochastic transition rules of the three strategies are:

- S^{ALLC} : $((1-\epsilon), (1-\epsilon), (1-\epsilon), (1-\epsilon))$,
- S^{ALLD} : $(\epsilon, \epsilon, \epsilon, \epsilon)$, and
- S^{Pavlov} : $((1-\delta)(1-\epsilon)(1-\zeta)+\delta\epsilon(1-\zeta)+\zeta\delta(1-\epsilon)+\zeta\epsilon(1-\delta), \epsilon(1-\delta)(1-\zeta)+\zeta(1-\delta)(1-\epsilon)+\delta(1-\epsilon)(1-\zeta), \epsilon(1-\delta)(1-\zeta)+\zeta(1-\delta)(1-\epsilon)+\delta(1-\epsilon)(1-\zeta), (1-\delta)(1-\epsilon)(1-\zeta)+\delta\epsilon(1-\zeta)+\zeta\delta(1-\epsilon)+\zeta\epsilon(1-\delta))$.

Note that neither ALLC nor ALLD is affected by perception errors of one's own or the opponent's action. Pavlov, however, is affected by all three types of errors.

A rule $\mathbf{p} = (p_1, p_2, p_3, p_4)$ that is matched against a rule $\mathbf{q} = (q_1, q_2, q_3, q_4)$ yields a Markov process where the transitions between the four possible states are given by the matrix

$$\begin{pmatrix} p_1q_1 & p_1(1-q_1) & (1-p_1)q_1 & (1-p_1)(1-q_1) \\ p_2q_3 & p_2(1-q_3) & (1-p_2)q_3 & (1-p_2)(1-q_3) \\ p_3q_2 & p_3(1-q_2) & (1-p_3)q_2 & (1-p_3)(1-q_2) \\ p_4q_4 & p_4(1-q_4) & (1-p_4)q_4 & (1-p_4)(1-q_4) \end{pmatrix}$$

If $\mathbf{p}$ and $\mathbf{q}$ are in the interior of the strategy cube, then all entries of this stochastic matrix are strictly positive. Consequently, there exists a unique stationary distribution $\pi^{\mathbf{p}/\mathbf{q}} = (\pi_1, \pi_2, \pi_3, \pi_4)$ such that $p_j^{(n)}$ is the probability of being in state j in the n^{th} period, and converges to π_j for $n \to \infty$ ($j = 1, 2, 3, 4$). It follows that the payoff for agent i using $\mathbf{p}$ against agent $-i$ using $\mathbf{q}$ is given by

$$\mathcal{P}(\mathbf{p}, \mathbf{q}) = R\pi_1 + S\pi_2 + T\pi_3 + P\pi_4, \qquad (1)$$

where the coefficients arise from the payoff matrix in Table 1. Notice that π_j and also the payoffs are independent of the initial condition. For any error level $0 < \epsilon, \delta, \zeta < 1$, the payoff obtained by a strategy using a transition rule S^i against a strategy with transition rule S^{-i} can be computed via (1). The limit value of the payoff for $\epsilon \to 0$, $\delta \to 0$, and $\zeta \to 0$ cannot be computed, as the transition matrix is no longer irreducible. Therefore, the stationary distribution π is no longer uniquely defined. In Table 2, we provide the asymptotic payoff matrix for any error level $0 < \epsilon, \delta, \zeta < 1$ across the three strategies: ALLD, ALLC, and Pavlov.

3 Axelrod's tournaments

There are many conceivable strategies for the repeated Prisoner's Dilemma game. This prompted Robert Axelrod to conduct computational tournaments to determine the best strategy in the game (Axelrod and Hamilton;[14] Axelrod[11]). In the first tournament, there were 14 entries, whereas in the second tournament, there were 63 entries. Tit-For-Tat (TFT) was the champion in both tournaments. TFT is a simple strategy that starts off by cooperating and then imitates the opponent's most recent action (Rapoport and Chammah[15]). TFT's victory was a startling outcome especially given that Axelrod had circulated the results and solicited entries from the first tournament prior to conducting the second tournament. Contestants in the second tournament tried to design more sophisticated strategies that

Table 2: Asymptotic Payoff Matrix in the Presence of Errors

Panel A	**ALLD**	**ALLC**	**Pavlov**
ALLD	$R\epsilon^2 + S\epsilon(1-\epsilon) +$ $T\epsilon(1-\epsilon) + P(1-\epsilon)^2$	$R(1-\epsilon)\epsilon + S\epsilon^2 +$ $T(1-\epsilon)^2 + P\epsilon(1-\epsilon)$	$[R(-\alpha\epsilon + \alpha\epsilon^2 - \beta\epsilon^2)\chi'(\epsilon-1) +$ $S(-\epsilon)\phi'\chi' + T(\alpha\epsilon - \alpha - \beta\epsilon)(\epsilon-1)\phi' +$ $P\phi'\chi'(\epsilon-1)]/\psi'$
ALLC	$R(1-\epsilon)\epsilon + S(1-\epsilon)^2 +$ $T\epsilon^2 + P\epsilon(1-\epsilon)$	$R(1-\epsilon)^2 + S\epsilon(1-\epsilon) +$ $T\epsilon(1-\epsilon) + P\epsilon^2$	$[R(-\alpha(1-\epsilon) + \alpha(1-\epsilon)^2 - \beta(1-\epsilon)^2)\chi''(-\epsilon) +$ $S(\epsilon-1)\phi''\chi'' + T(\alpha(1-\epsilon) - \alpha - \beta(1-\epsilon))\phi''(-\epsilon) +$ $P\phi''\chi''(-\epsilon)]/\psi''$
Pavlov	$[R(-\alpha\epsilon + \alpha\epsilon^2 - \beta\epsilon^2)\chi'(\epsilon-1) +$ $S(\alpha\epsilon - \alpha - \beta\epsilon)(\epsilon-1)\phi' + T(-\epsilon)\phi'\chi' +$ $P\phi'\chi'(\epsilon-1)]/\psi'$	$[R(-\alpha(1-\epsilon) + \alpha(1-\epsilon)^2 - \beta(1-\epsilon)^2)\chi''(-\epsilon) +$ $S(\alpha(1-\epsilon) - \alpha - \beta(1-\epsilon))\phi''(-\epsilon) +$ $T(\epsilon-1)\phi''\chi'' + P\phi''\chi''(-\epsilon)]/\psi''$	$[R(\alpha^2\beta - \alpha\beta^2 - \frac{\alpha^2}{2})\chi +$ $S\alpha\phi + T\alpha\phi + P\phi\chi]/\psi$
Panel B			
	$\alpha = (1-\delta)(1-\epsilon)(1-\zeta) + \delta\epsilon(1-\zeta) + \zeta\delta(1-\epsilon) + \zeta\epsilon(1-\delta)$		
	$\beta = \epsilon(1-\delta)(1-\zeta) + \zeta(1-\delta)(1-\epsilon) + \delta(1-\epsilon)(1-\zeta)$		
	$\phi = -\alpha^2\beta + \frac{\alpha^2}{2} + \alpha\beta^2 - \beta^2 + \beta - \frac{1}{2}$		
	$\chi = -2\alpha\beta + \alpha + 2\beta^2 - 2\beta + 1$		
	$\psi = -\frac{1}{2}\left(-\alpha + 2\beta - 2\beta^2 + 2\alpha\beta - 1\right)\left(-2\alpha + 2\beta + 2\alpha^2 - 2\beta^2 - 1\right)$		
	$\phi' = \beta + \epsilon + \alpha\epsilon - 2\beta\epsilon - \alpha\epsilon^2 + \beta\epsilon^2 - 1$		
	$\chi' = \beta + \alpha\epsilon - \beta\epsilon - 1$		
	$\psi' = (\epsilon - 1)\left(-\alpha + \beta + 2\alpha\epsilon - 2\beta\epsilon - 1\right)\left(\beta + \alpha\epsilon - \beta\epsilon - 1\right)$		
	$\phi'' = \beta + (1-\epsilon) + \alpha(1-\epsilon) - 2\beta(1-\epsilon) - \alpha(1-\epsilon)^2 + \beta(1-\epsilon)^2 - 1$		
	$\chi'' = \beta + \alpha(1-\epsilon) - \beta(1-\epsilon) - 1$		
	$\psi'' = \epsilon\left(\alpha - \alpha\epsilon + \beta\epsilon - 1\right)\left(-\alpha + \beta + 2\alpha\epsilon - 2\beta\epsilon + 1\right)$		

Notes: In *Panel A*, we provide the asymptotic payoff matrix for any error level $0 < \epsilon, \delta, \zeta < 1$ across the three strategies: ALLD, ALLC, and Pavlov. The successive actions chosen by agents using such strategies are modeled by a Markov process. The presence of errors guarantees that the process is ergodic. Computation of the invariant distribution of the realized actions for each pair of strategies, yields the payoffs in the matrix. The payoffs are those of the row agent when paired with a column agent and utilize the letters in the generic Prisoner's Dilemma matrix in Table 1. In *Panel B*, we provide the expressions that have been replaced by the Greek letters in *Panel A*.

were superior to TFT, yet TFT won again (Axelord;[16] Axelord[17]). The key to TFT's success is that it cooperates with other reciprocators, but resists exploitation by strategies, such as ALLD. Nevertheless, TFT has an Achilles' heel which did not become apparent in the error-free tournaments of Axelrod (Nowak and Sigmund[18]). If a pair of TFTs interact with each other and one defects by mistake, then the other TFT will retaliate and thus the two TFTs will lock themselves into an endless string of retaliations (i.e. a vendetta). TFT's unrelenting punishment never forgives even a single deviation, thus inhibiting the evolution of cooperation (Molander[3]).

Following the success of Axelrod's tournaments, Bendon et al.[19] initiated a new tournament. However, in this tournament the authors re-evaluated the performance of reciprocating strategies, such as TFT, and identified alternative strategies that could sustain cooperation in an environment with random shocks. This time TFT placed eighth out of the thirteen strategies considered. The winning strategy was Nice-And-Forgiving (NAF), which differs in many ways from TFT. First, NAF is nice in the sense that it cooperates as long as the frequency of cooperation of the opponent is above some threshold. Second, NAF is forgiving in the sense that, although NAF retaliates if the opponent's cooperation falls below the threshold level, it reverts to full cooperation before its opponent does, as long as the opponent meets certain minimal levels of cooperation. The success of NAF is not a robust result, but is limited to the particular environment. As Bendor et al.[19] note, the generosity of NAF creates a risk: other strategies may exploit NAF's willingness to give more than it receives. In other words, NAF can be suckered by a nasty strategy that is disinterested in joint gains. Due to its generosity, NAF lost in its pairwise play with every one of its opponents. In contrast to NAF's pattern, VIGILANT, the strategy that placed dead last in the tournament, beat every one of its partners in bilateral play. VIGILANT was a highly provocative and unforgiving strategy that retaliated sharply if it inferred that its partner was playing anything less than maximal cooperation. Beyond NAF, Nowak and Sigmund[20] showed that Pavlov could also typically outperform TFT in the repeated Prisoner's Dilemma game. Pavlov responds to stimuli applying Thorndike's "law of effect," hence it is more resistant than TFT to cycles of recrimination and thrives at the expense of unconditional cooperative strategies. Indeed, if a pair of Pavlovs interact with each other and one defects in error, then the other Pavlov will retaliate in the next period (so will the Pavlov that committed the error), but subsequently both Pavlovs will resume mutual cooperation. In addition, if Pavlov notices than an inadvertent defection against an opponent meets no retaliation, Pavlov will continue to defect thus capitalizing on the "Temptation" payoff. Thus, Pavlov would exploit a strategy, such as ALLC. Yet Pavlov is vulnerable to exploitation by unconditionally aggressive strategies, such as ALLD. On the other hand, ALLD is myopically flawed despite being aggressive as its short-term advantage succumbs to its own success.

Within the huge class of strategies in the repeated Prisoner's

Dilemma game, we shall concentrate on three, memory-one strategies: ALLD, ALLC, and Pavlov. ALLD and ALLC are natural choices to provide a lower and an upper bound on payoffs, and Pavlov has been shown to be evolutionary stable in the set of strategies of finite complexity when allowing for errors (Fudenberg and Maskin[4]). Needless to say, there are many other possible strategies that could be considered, and some indeed play an important role. Yet coping with a mere, tri-morphic population still results in a high degree of perplexity in terms of selection dynamics, mainly because of the possibility of "rock-paper-scissors" cycles; that is, it may happen that strategy B beats strategy A, strategy C beats strategy B, and strategy A, in turn, beats strategy C. In such a case, the selection dynamics can lead to long-term coexistence of the strategies.

4 Numerical results

In this section, we first determine how the level of errors affects the payoffs of a pair of agents. We then display graphically the payoffs of the strategies in the head-to-head competitions for the entire range of errors. We also present examples to discuss the implications to the population dynamics, and provide a taxonomy of best response strategies in the presence of errors. For the calculation of the values, we used the payoff matrix from the tournaments of Axelrod.[1] The payoff matrix is displayed in Table 3. Our quest to derive exact results led us to impose two simplifying assumptions. We assume that (i) the error level is *common knowledge* amongst agents, and (ii) for each error type, the error level is the same. To fix ideas, consider a parasite's virulence, which depends only on the neighborhood. Furthermore, assume that all adjacent parasites face the exact identical conditions − the only difference amongst the adjacent parasites is the strategy pursued (Nowak and May[21]).

Table 3: Axelrod's Prisoner's Dilemma Payoff Matrix

	Cooperate	**Defect**
Cooperate	3	0
Defect	5	1

Note. The payoffs are those of the row agent

4.1 Payoffs with errors

As highlighted above, we shall assume that the error level is common knowledge and that the likelihood of committing any type of error is the same; that is, $0 < \epsilon = \delta = \zeta < 1$ for all levels. Note that the latter assumption is not as restrictive as it may seem. Recall that only Pavlov is affected by all three types of errors; ALLD and ALLC are only affected by implementation errors. Having said this, we do acknowledge that varying the error levels across error types would contribute to the generality of the results. We defer such an interesting venue for future work.

The idea is to first calculate the invariant distribution of each pair combination at each error level. Then, the distribution is plugged into the payoff function (1) with the numerical values of Axelrod's payoff matrix in order to derive the corresponding payoffs. For example, assume that an ALLD is matched with another ALLD at an error level of 10%. The invariant distribution at this error level is $\pi^{ALLD/ALLD} = (0.10^2, 0.10 \cdot (1-0.10), 0.10 \cdot (1-0.10), (1-0.10)^2)$. The payoff for agent i using ALLD against agent $-i$ also using ALLD, at the 10% error level, is given by $\mathcal{P}(ALLD, ALLD) = 3 \cdot 0.10^2 + 0 \cdot 0.10 \cdot (1-0.10) + 5 \cdot 0.10 \cdot (1-0.10) + 1 \cdot (1-0.10)^2 = 1.29$. The same procedure is followed for all combination pairs in all error levels. This way, we can observe the trends, as well as any monotonicity properties where those exist. The payoffs for the entire range of errors is demonstrated graphically in Figure 1. It is interesting to note that with the exception of the payoff of an ALLC when paired with a Pavlov, which exhibits no monotonicity properties, every other payoff function is either monotonically increasing or decreasing as the likelihood of errors increases.

In the top left corner in Figure 1, the plot of the payoff of ALLC when paired with one of ALLC, ALLD, or Pavlov is provided. The payoff of an ALLC paired with a twin starts off near the "Reward" payoff, but goes down and reaches near the "Punishment" payoff as the likelihood of committing implementation errors reaches 1. In the specific mark, both ALLCs behave almost as ALLDs. The payoff of an ALLC when matched with an ALLD starts off near the "Sucker's" payoff and gradually comes near to the "Temptation" payoff when the likelihood of committing an implementation error approaches 1. At this mark, ALLC acts almost as if it were an ALLD, whereas an ALLD acts almost as if it were an ALLC. On the other hand, the payoff of an ALLC when paired with Pavlov starts off a little bit higher than the "Punishment" payoff, then at the 50% error level comes close to a payoff of 2 as the pair alternates between the "Reward" payoff and the "Punishment" payoff. Around the 55% error level, the payoff of ALLC reverses direction. At such high error levels, ALLC acts as an ALLD in low error levels. Thus, the outcomes realized by a pair of an ALLC and a Pavlov, at high error levels, tend to be more defecting than cooperative.

In the top right corner in Figure 1, the plot of the payoff of an ALLD when matched to one of the three strategists is displayed. The payoff of an ALLD paired with an ALLC starts off near the "Temptation" payoff, but as the probability of committing implementation errors approaches 1, the payoff of an ALLD draws near the "Sucker's" payoff. At this specific mark, as explained earlier, there is a reversal of roles; that is, an ALLD acts almost as if it were an ALLC, whereas an ALLC acts almost as if it were an ALLD. The payoff of an ALLD paired with a twin starts off near the payoff of 1 and draws closer to a payoff of 3 as the likelihood of implementation errors approaches 1. At this specific mark, both ALLDs behave almost as ALLCs. The payoff of an ALLD when paired with Pavlov starts off near a payoff of 3 and gradually decreases close to the "Sucker's" payoff when the likelihood of committing errors approaches 1.

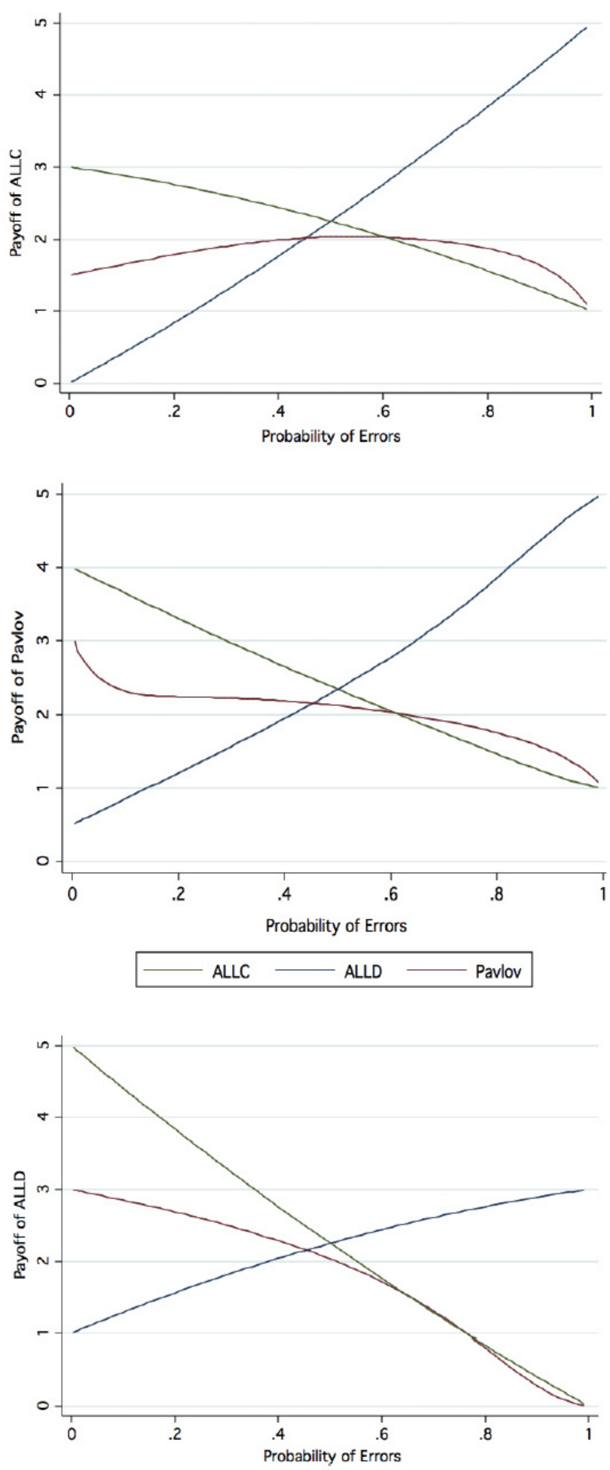

Figure 1: Payoffs in the Presence of Errors

In the bottom left corner in Figure 1, the plot of the payoff of a Pavlov when paired with one of ALLC, ALLD, or Pavlov is presented. The payoff of a Pavlov when paired with an ALLC starts off close to midway of the "Temptation" and "Reward" payoffs, before getting closer to the "Punishment" payoff as the likelihood of committing errors approaches 1. The payoff of a Pavlov when matched with an ALLD starts off a little bit higher than the "Sucker's" payoff and gradually comes near to the "Temptation" payoff when the likelihood of committing errors reaches 1. The payoff of a Pavlov paired with a twin starts off near the "Reward" payoff, but moves closer the "Punishment" payoff as the likelihood of committing errors approaches 1.

4.2 Taxonomy of best response strategies in the presence of errors

In the previous subsection, we plotted the payoffs of the strategies in the head-to-head competitions for the entire range of error levels. Here, we look at one error level each time to determine the best response strategy for the error level specified. When a *unique* best response is indicated, then that best response is also a dominant strategy. The notion of dominance is significant to the selection dynamics. More specifically, an ecological selection process will enable the dominant strategy to proliferate to the point where the entire population converges to a pure one implementing that strategy. Thus, depending on the error level, an evolving population could consist of only ALLDs, only ALLCs, or only Pavlovs. We present next, examples and discuss the implication for the population dynamics. We then provide a taxonomy with the best response strategies in the presence of errors.

- *Example 1 ($\epsilon = \delta = \zeta = 20\%$)*

Table 4: Payoff Matrix With $\epsilon = \delta = \zeta = 20\%$.

	ALLD	ALLC	Pavlov
ALLD	1.56	3.84	2.69
ALLC	0.84	2.76	1.79
Pavlov	1.20	3.31	2.24

Note. Tables 4-6 follow the same structure. The payoff matrix indicates the payoffs of a population consisting of ALLDs, ALLCs and Pavlovs. The payoffs are those of the row agent when paired with a column agent and reflect the invariant distributions for the error level specied. In this example, the dominant strategy is ALLD.

In this first example, we assume that implementation errors, perception errors of the opponent's action, and perception errors of one's own action are kept constant at 20%. The payoff matrix of a population consisting of ALLDs, ALLCs, and Pavlovs is indicated in Table 4. Clearly, this is a dominance solvable game, where ALLD is the dominant strategy. Therefore, based on the specific error level, an ecological selection process would converge to a population consisting of only ALLDs.

On the other hand, a different outcome would emerge for an error level of 50%. In this case, the process would converge to a population consisting of only Pavlovs.

- *Example 2* ($\epsilon = \delta = \zeta = 50\%$)

The payoff matrix in Table 5 corresponds to implementation errors, perception errors of the opponent's action, and perception errors of one's own action kept constant at 50%. Analogous to the previous example, there is a dominant strategy – Pavlov. Note that an ALLD and an ALLC earn identical payoffs as, at this error level, the two strategies emulate the strategy Random. On the other hand, Pavlov is a somewhat less cooperative form of Random. Kraines and Kraines[9] show that Random is a poor strategy against strategies that defect frequently. This result is also demonstrated here.

Table 5: Payoff Matrix With $\epsilon = \delta = \zeta = 50\%$.

	ALLD	ALLC	Pavlov
ALLD	2.25	2.25	2.03
ALLC	2.25	2.25	2.03
Pavlov	2.34	2.34	2.12

Note. The dominant strategy is Pavlov.

The selection dynamics therefore depend crucially on the error levels. Figure 2 is indicative of the error level dependence. The horizontal axis indicates the error level. The bar indicates any strategy that is a best response to another strategy. Up to the 45.3% error level, ALLD is the dominant strategy, as it is the best response to each of the strategies. Between error levels 45.4% and 60.4%, the dominant strategy is Pavlov. Furthermore, from error level 64.2% to 76%, the dominant strategy is ALLC.

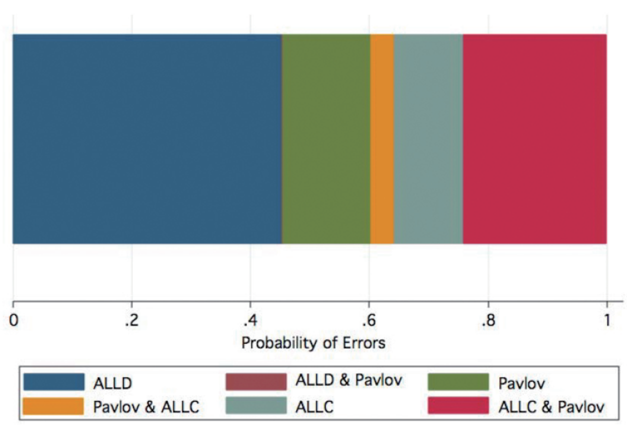

Figure 2: Taxonomy of Best Response Strategies in the Presence of Errors

A natural question to ask is what will be the outcome of an ecological process when a unique best response does not exist. In other words, what would the composition of the population be like in the range of error levels *in transit* from one dominant strategy to another. The population composition could vary based on the proportion of each strategist's type and/or the payoffs of the Prisoner's Dilemma matrix. We illustrate these features of the population dynamics with an example.

- *Example 3* ($\epsilon = \delta = \zeta = 62\%$)

In this last example, we assume that implementation errors, perception errors of the opponent's action, and perception errors of one's own action are kept constant at 62%. The payoff matrix is indicated in Table 6. In contrast to the two previous examples, there is no dominant strategy. ALLC is a best response if the opponent uses ALLC or Pavlov, but Pavlov is a best response if the opponent uses ALLD. Interestingly, the population dynamics depend on the proportion within the population of each type and the payoff choices. For expositional purposes, let us fix the population size at 30 agents. The agents can use any one of the three strategies, and are paired in a round-robin structure; that is, all agents are paired with one another in every possible combination. Furthermore, we assume that a strategy does *not* play itself. To calculate the average payoff of each pair, we use the values of the payoff matrix in Table 3. Finally, assume that within the population of 30 agents, x agents use ALLC, y agents use ALLD, where $x, y \in N, N = \{1, 2, ..., 29\})$. In particular, the payoffs, at this error level, are the following:

$$P(\mathrm{A}LLD) = \frac{2.48 \cdot (y-1) + 1.66 \cdot x + 1.64 \cdot (30-x-y)}{29},$$

$$P(\mathrm{A}LLC) = \frac{2.00 \cdot (x-1) + 2.02 \cdot (30-x-y) + 2.86 \cdot y}{29},$$

$$P(\mathrm{P}avlov) = \frac{1.99 \cdot (x) + 2.01 \cdot (30-x-y-1) + 2.87 \cdot y}{29}.$$

If $x = 1$, $y = 19$, and 10 play Pavlov, then, at this error level, $P(\mathrm{P}avlov) > P(\mathrm{A}LLC) > P(\mathrm{A}LLD)$. In a more balanced population though, where $x = 10$, $y = 10$, and 10 play Pavlov, then $P(\mathrm{A}LLC) > P(\mathrm{P}avlov) > P(\mathrm{A}LLD)$.

Table 6: Payoff Matrix With $\epsilon = \delta = \zeta = 62\%$.

	ALLD	ALLC	Pavlov
ALLD	2.48	1.66	1.64
ALLC	2.86	2.00	2.02
Pavlov	2.87	1.99	2.01

Note. There is no dominant strategy.

However, to illustrate the sensitivity of population dynamics to the choice of payoffs in the matrix, we alter the "Temptation" payoff and the "Reward" payoff of Table 3. In particular, the "Temptation" payoff is changed to 11 from 5, and the

"Reward" payoff is changed to 6 from 3. The "Punishment" payoff and the "Sucker's" payoff stay the same. Analogous to the initial case, assume that $x = 1$ agent uses ALLC, $y = 19$ agents use ALLD, and $30 - 1 - 19 = 10$ use Pavlov. The payoffs, for an error level kept constant at 62%, are the following:

$$P(\mathrm{A}LLD) = \frac{5.04\cdot(y-1)+3.24\cdot x+3.19\cdot(30-x-y)}{29},$$

$$P(\mathrm{A}LLC) = \frac{3.84\cdot(x-1)+3.90\cdot(30-x-y)+5.88\cdot y}{29},$$

$$P(\mathrm{P}avlov) = \frac{3.82\cdot(x)+3.86\cdot(30-x-y-1)+5.90\cdot y}{29}.$$

Therefore, $P(\mathrm{A}LLC) > P(\mathrm{P}avlov) > P(\mathrm{A}LLD)$ in sharp contrast to the former ranking $P(\mathrm{P}avlov) > P(\mathrm{A}LLC) > P(\mathrm{A}LLD)$, which was based on the payoff matrix in Table 3.

5 Discussion

TFT was the winner in the *in silico* tournaments of Axelrod. The performance of TFT led Axelrod to identify four basic attributes that were necessary for the emergence and survival of cooperation: (i) provocation in the face of an uncalled-for defection by the other; (ii) forgiveness after responding to a provocation; (iii) clarity of behavior so that the other agent can adapt to your pattern of action; and (iv) avoidance of unnecessary conflict by cooperating as long as the other agent cooperates. Yet Axelrod and Dion[22] recognized that in the presence of errors, the emergence of cooperation is hardly inevitable; unnecessary conflict can only be avoided by generosity, but generosity invites exploitation, and exploitation invites retaliation. Dawes and Thaler[23] point out that many *in vivo* experimental studies of the repeated Prisoner's Dilemma game, have found that mere suspicion of the possibility of exploitation induces individuals to engage in a kind of defensive "stinginess" that imparts the emergence of cooperation. More recently, Dal Bó and Fréchette[24] provide compelling experimental evidence with human data to suggest that even in treatments where cooperation can be supported in equilibrium, the level of cooperation may remain at low levels even after significant experience is obtained. The authors conclude that "these results cast doubt on the common assumption that agents will make the most of the opportunity to cooperate whenever it is possible to do so in equilibrium" (p. 412).

In this study, we consider three simple, but fundamental strategies in the presence of the entire range of errors. Our findings confirm that the emergence of cooperation is not as likely at any error level. In the error levels where a dominant strategy exists, ALLD is the dominant strategy at low error levels; Pavlov is the dominant strategy at intermediate error levels; and ALLC, which acts as if it were an ALLD in low error levels, is the dominant strategy at high error levels. In almost all of these cases, non-cooperative outcomes emerge; that is, vulnerability to errors prompts non-cooperative outcomes and, oftentimes, results in long strings of retaliations (i.e. vendettas).

Vulnerability to errors is a real concern in international politics, and has been an especially poignant one during the Cold War. Speaking of the likelihood of nuclear accidents in the wake of the Cuban Missile Crisis in 1962, Assistant Secretary of Defense John T. McNaughton stated that, "the explosion of a nuclear device by accident − mechanical or human − could be a disaster for the United States, for its allies, and for its enemies. If one of these devices accidentally exploded, I would hope that both sides had sufficient means of verification and control to prevent the accident from triggering a nuclear exchange. But we cannot be certain that this would be the case" (Sagan[25]). United States was indeed concerned by accidents, such as the B-52 crash in 1968, near Thule Air Base in Greenland. Along with the U.S.S.R., they agreed to take measures to ensure that a future nuclear accident would not lead the other party to conclude incorrectly that a first strike was under way. Consequently, on 30 September, 1971, the two superpowers signed the "Agreement on Measures to Reduce the Risk of Nuclear War." Each party agreed to notify the other immediately in the event of an accidental, unauthorized or unexplained incident involving a nuclear weapon that could increase the risk of nuclear war. They even agreed to use the Moscow-Washington hotline, which was upgraded at the same time, for any communication between the two countries. Also, following the downing of Korean Airlines Flight 007 after it strayed over territory belonging to the Soviet Union, Reagan announced the expansion of the Global Positioning System (GPS) to civilians, which at the time was only used by the US military. It would thus be harder for any pilots to drift into Soviet airspace with satellite navigation technology. In 1987, the Department of Defense formally requested the Department of Transportation to establish and provide an office to respond to civil users' needs and to work closely with the Department of Defense to ensure proper implementation of GPS for civil use.

Finally, we find that in the error levels where a dominant strategy does not exist, the population composition can vary based on the proportion of each strategist's type and/or the payoffs of the Prisoner's Dilemma matrix. Several studies have considered variations in the Prisoner's Dilemma payoff matrix, most often to examine its effect on the likelihood of cooperation (Busch and Reinhardt;[26] Stephens et al.[27]). Here, we go one step further, to show that, not only, the payoff structure affects selection dynamics and hence the likelihood of cooperation, but also the proportion of each strategist's type. Consequently, in such cases it becomes possible to observe "rock-paper-scissors" cycles, where the selection dynamics can lead to long-term coexistence of the strategies. A lot of examples of long-term co-existence can be found in nature. For instance, there exist three morphs of the male territorial iguanid lizard *Uta stansburiana* who differ in their throat color and in their mate-guarding behavior. Males with orange throats are monogamous and succeed in preventing other males from approaching their mates. Males with dark blue throats are polygamous and less efficient, having to split their efforts on several females. Males with prominent yellow stripes on their throats do not engage in female-guarding behavior at all, but roam around in search of sneaky matings (Sinervo and Lively[28]). Furthermore, there exist three strains of *Escherichia coli* bacteria. The colicin-producing strain releases toxic colicin and produces, for its own protection, an immunity protein. The sensitive strain produces the immunity protein only. The resistant strain, on the other hand, produces neither the toxic colicin nor the immunity protein (Kerr et al.[29]).

6 Concluding remarks

We study the asymptotic behavior of a finite, but error-prone population of agents that play the repeated Prisoner's Dilemma game. Three types of errors are allowed: implementation errors, perception errors of one's own action, and perception errors of the opponent's action. The space of strategies considered is restricted to simplify the dynamics. We thus concentrate on an ALLD strategist, an ALLC strategist, and a Pavlov strategist. We first obtain the asymptotic payoff matrix across the three strategies in the presence of errors, and then derive numerical results using the payoff matrix from Axelrod's celebrated tournaments. Crucially, we consider the entire range of error levels in order to classify best response strategies as the likelihood of error levels increases from zero to one. We find that for some range of error levels, a unique best response (a dominant strategy) exists. Moreover, the dominant strategies at the corresponding error levels lead to mostly non-cooperative outcomes. On the other hand, in the range of error levels where a dominant strategy does not exist, the population composition can vary based on the proportion of each strategist's type and/or the payoffs of the Prisoner's Dilemma matrix. In such cases, it is possible to observe "rock-paper-scissors" cycles.

Our findings highlight that a systematic analysis of the entire range of error levels is an important and essential aspect of population dynamics. Thus, an interesting direction for future research would be to use the methodology prescribed to provide a taxonomy of best response strategies in the presence of errors in alternative games. Ultimately, one would like to determine better strategies across a vast array of games and for different levels of errors. Another promising direction for future research would be to assume that there are differences in the levels of errors across the three error types. For instance, there exist environments where agents are more likely to commit perception errors of the opponent's action (perhaps, due to limited channels in the transmission of information) than implementation errors.

References

[1] Robert Axelrod. The Evolution of Cooperation. Basic Books: New York, 1984.

[2] Reinhard Selten. A re-examination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory. 1975; 4: 25-55. http://dx.doi.org/10.1007/BF01766400

[3] Per Molander. The optimal level of generosity in a selfinsh, uncertain environment. Journal of Conict Resolution. 1985; 29: 611-8. http://dx.doi.org/10.1177/0022002785029004004

[4] Drew Fudenberg and Eric Maskin. Evolution and cooperation in noisy repeated games. American Economic Review. 1990; 80: 274-9.

[5] Martin A. Nowak, Karl Sigmund, and Esam El-Sedy. Automata, repeated games, and noise. Journal of Mathematical Biology. 1995; 33: 703-22.

[6] David P. Kraines and Vivian Y. Kraines. Evolution of learning among pavlov strategies in a competitive environment with noise. Journal of Conict Resolution. 1995; 39: 439-66. http://dx.doi.org/10.1177/0022002795039003003

[7] Jianzhong Wu and Robert Axelrod. How to cope with noise in the iterated prisoner's dilemma. Journal of Conict Resolution. 1995; 39: 183-89. http://dx.doi.org/10.1177/0022002795039001008

[8] M. Lindi Wahl and Martin A. Nowak. The continuous prisoner's dilemma: linear reactive strategies with noise. Journal of Theoretical Biology. 1999; 200: 323-38. PMid:10527721 http://dx.doi.org/10.1006/jtbi.1999.0997

[9] David P. Kraines and Vivian Y. Kraines. Natural selection of memory-one strategies for the iterated prisoner's dilemma game. Journal of Theoretical Biology. 2000; 203: 335-355. PMid:10736212 http://dx.doi.org/10.1006/jtbi.2000.1089

[10] Karthik Panchanathan and Robert Boyd. A tale of two defectors: The importance of standing for evolution of indirect reciprocity. Journal of Theoretical Biology. 2003; 224: 115-26. http://dx.doi.org/10.1016/S0022-5193(03)00154-1

[11] Martin A. Nowak and Karl Sigmund. Evolutionary dynamics of biological games. Science. 2004; 303: 793-9. PMid:14764867 http://dx.doi.org/10.1126/science.1093411

[12] Lorens A. Imhof, Drew Fudenberg, and Martin A. Nowak. Evolutionary cycles of cooperation and defection. Proceedings of the National Academy of Science. 2005; 102(31): 797-800. PMid:15642942

[13] David G. Rand, Drew Fudenberg, and Anna Dreber. It's the thought that counts: The role of intentions in reciprocal altruism. Working Paper, 2013.

[14] Robert Axelrod and William D. Hamilton. The evolution of cooperation. Science. 1981; 211:1390-6. http://dx.doi.org/10.1126/science.7466396

[15] Anatol Rapoport and Albert M. Chammah. Prisoner's Dilemma. Ann Arbor: University of Michigan Press, 1965.

[16] Robert Axelrod. Effective choice in the prisoner's dilemma. Journal of Conict Resolution. 1980; 24: 3-25.

[17] Robert Axelrod. More effective choice in the prisoner's dilemma. Journal of Conict Resolution. 1980; 24: 379-403. http://dx.doi.org/10.1177/002200278002400301

[18] Martin A. Nowak and Karl Sigmund. Tit-for-tat in heterogeneous populations. Nature. 1992; 355: 250-3. http://dx.doi.org/10.1038/355250a0

[19] Jonathan Bendor, Roderick M. Kramer, and Suzanne Stout. When in doubt cooperation in a noisy prisoner's dilemma. Journal of Conict Resolution. 1991; 35(4): 691-719. http://dx.doi.org/10.1177/0022002791035004007

[20] Martin A. Nowak and Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in prisoner's dilemma. Nature. 1993; 364: 56-8. PMid:8316296 http://dx.doi.org/10.1038/364056a0

[21] Martin A. Nowak and Robert M. May. Virus Dynamics. Oxford Press, 2000.

[22] Robert Axelrod and Douglas Dion. The further evolution of cooperation. Science. 1988; 242: 1385-90. PMid:17802133 http://dx.doi.org/10.1126/science.242.4884.1385

[23] Robyn M. Dawes and Richard H. Thaler. Anomalies: Cooperation. Journal of Economic Perspectives. 1988; 2: 187-97. http://dx.doi.org/10.1257/jep.2.3.187

[24] Pedro Dal Bo and Guillaume R. Frechette. The evolution of cooperation in innitely repeated games: Experimental evidence. American Economic Review. 2011; 101(1): 411-29. http://dx.doi.org/10.1257/aer.101.1.411

[25] Scott Sagan. The Limits of Safety: Organizations, Accidents, and Nuclear Weapons. Princeton, NJ: Princeton University Press, 1995.

[26] Marc L. Busch and Eric Reinhardt. Nice strategies in a world of relative gains: The problem of cooperation under anarchy. Journal of Conict Resolution. 1993; 37: 427-45. http://dx.doi.org/10.1177/0022002793037003002

[27] David W. Stephens, Kinya Nishimura, and K. B. Toyer. Error and discounting in the iterated prisoner's dilemma. Journal of Theoretical Biology. 1995; 176(4): 457-69. http://dx.doi.org/10.1006/jtbi.1995.0213

[28] Barry Sinervo and Curt M. Lively. The rock-paper-scissors game and the evolution of alternative male strategies. Nature. 1996; 380: 240-3. http://dx.doi.org/10.1038/380240a0

[29] Benjamin Kerr, Margaret A. Riley, Marcus W. Feldman, and Brendan J. M. Bohannan. Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. Nature. 2002; 418: 171-4. PMid:12110887 http://dx.doi.org/10.1038/nature00823