Equilibria analysis in social dilemma games with Skinnerian agents

Published version:

DOI:10.1007/s11299-013-0116-6

Ugo Merlone;Daren R. Sandbank;Ferenc Szidarovszky. Equilibria analysis in social dilemma games with Skinnerian agents. MIND & SOCIETY. 12 (2) pp: 219-233.
DOI: 10.1007/s11299-013-0116-6

Equilibria analysis in social dilemma games with Skinnerian agents

Ugo Merlone · Daren R. Sandbank · Ferenc Szidarovszky

Abstract Different disciplines have analyzed binary choices to model collective behavior in human systems. Several situations in which social dilemma arise can be modeled as *N*-person prisoner's dilemma games including homeland security, public goods, international political economy among others. The purpose of this study is to develop an analytical solution to the *N*-person prisoner's dilemma game when boundedly rational agents interact in a population. Previous studies in the literature consider the case in which cooperators and defectors have the same learning factors. We obtain an analytical equation to find equilibria in the *N*-person prisoner's dilemma game in the general case when agents have different learning factors. We also introduce a more realistic approach where probability values are bounded between zero and one and therefore eliminates the possibility of infeasible probability values. Since no analytic solution can be derived in this case, agent based simulation is used to analyze the asymptotic behavior of the resulted dynamical system.

Keywords Social dilemmas · Binary games · Bounded rationality · Agent based simulation

1 Introduction

Social dilemmas are situations in which each individual has a clear and unambiguous incentive to make a choice that provides a poorer outcome for all when it is made by all individuals than the outcome they would have received when none of them had made the choice (Dawes and Messick 2000, p. 111).

Social dilemmas arise from collective actions in any part of or the entire society and have been discussed in several disciplines: economics (Ledyard 1995; Ostrom 1998), psychology (Dawes 1980), sociology (Kollock 1998). Furthermore, particular applications can be developed in homeland security (Brams and Kilgour 1988), public goods (Tabarrok 1998), international political economy (Conybeare 1984) among others.

According to Kollock (1998, p. 183), "social dilemmas are situations in which individual rationality leads to collective irrationality". Among other social dilemmas, N-person prisoner's dilemma game has come to be viewed as one of the most common representations of collective action problems (Ostrom 2000).

According to Santos et al. (2008, p. 213), the "N-person prisoners dilemma constitutes the most used metaphor to study public goods games", yet according to the literature the two games are not immediately equivalent. For example, Conybeare (1984) analyzes the differences between the prisoners' dilemma and public good games and a classification depending on the existence of provision point in the game is provided in Ledyard (1995). Nevertheless Hauert and Szabo (2003) generalize prisoner's dilemma to an arbitrary number of players and by a simple transformation link the resulting game to a public good game.

The prisoner's dilemma and more generally the N-person prisoner's dilemma can be considered a binary game with externalities (Schelling 1973). In the recent literature, this kind of interactions has been analyzed by several contributions. Bischi and Merlone (2009) provides a discrete-time dynamical system to model a class of binary choice games with externalities as those described by Schelling (1973). The dynamical properties of this game is studied in Bischi et al. (2009a, b); Gardini et al. (2011); Dal Forno et al. (2012) extendes the analysis to ternary choice games. Another intriguing social dilemma, the Braess paradox, is studied in Dal Forno and Merlone (2013). Finally, when considering binary games with interactions limited to neighborhoods, Merlone et al. (2007) proves that the number of different types of games is finite and then tight upper bounds for the number of game types were derived in the general case and the different game types were identified.

In the study of social dilemma learning has an important role as it may be a possible explanation for the emergence of norms (Ostrom 2000), the dynamics of cooperation, but also as a way to make agents to base their predictions on experiential induction rather than logical deduction (Macy and Flache 2002).

Other authors concentrated on the N-person prisoner's dilemma considering both behavior patterns to model boundedly rational agents and learning. In Szilagyi (2003) some interesting results show strong dependence on the choice of model parameter values. Nevertheless, in this contribution there are some strong assumptions about learning. In fact, Szilagyi (2003) considers only the case in which cooperators and defectors have the same learning factors and provides an algebraic equation to characterize equilibria.

Assuming that cooperators and defectors have the same learning factors is a serious limitation for several real life applications. For example, consider developing a homeland security model where a number of countries unite to fight terrorism activities from a specific group or threat. Each country agrees to contribute to the effort at a different level based on its size, resources, susceptibility to the threat and other factors. In this model, the payoff for each country is a level of security less its contribution to the effort. This is an N-person prisoners dilemma game where the learning curves for a cooperating and defecting country would likely be different because a defector would have to consider the negative political and foreign relation consequences in additional to the payoff when considering its next position.

The purpose of this study is to relax this assumption and to develop an analytical solution to the N-person prisoner's dilemma game when agents have different learning factors.

Our equation is a straightforward generalization of the equilibrium equation with equal learning factors. As it will be explained later, both equations hold only under certain conditions. We will also develop a modified dynamic model where these restrictions are eliminated. However, in this more realistic case no analytical solution is found. Therefore agent based simulation is used to examine the dynamic properties of the system.

In an agent based simulation model of binary games individual agents may cooperate with each other for the collective best interests or defect to pursue their own self interest. Each agent receives a reward or punishment based on its individual decision and the collective decisions of the other agents. The study of social dilemma problems is very important since it helps us understand problems that society is facing today. The prisoner's dilemma is an important example of a social dilemma which is frequently examined in the literature.

Agent based social simulation started in the 1990s, (Gilbert and Terna 2000) and this research area is growing significantly. The main advantage of agent based simulation is that it is a bottoms-up approach where the agents' attributes may include personalities, characteristics and learning capabilities. Agent based simulation can be used to study artificial societies for emergence of groups with common attributes or social segregation (Epstein and Axtell 1996).

This methodology is used in this paper for two purposes. First the analytical solution is verified when its conditions are satisfied, and second, in the more realistic model when no analytical solution is available, it is the most appropriate tool to analyze the dynamic properties of the model, because, as mentioned in Young (1998), simulation can be used to establish constructive sufficiency.

The structure of the paper is the following. In Sect. 2 we summarize the prisoner's dilemma and its N-person version which is the object of our analysis. In Sect. 3 we derive the analytic solution for Skinnerian agents. The cases of other types of boundedly rational agents will be discussed briefly. In Sect. 4 we describe the simulation analysis and compare the simulation and theoretical results. Finally, the last section is devoted to conclusions and further research directions.

2 The N-person prisoner dilemma

The prisoner's dilemma is typically defined as a two player discrete game (Axelrod 1984). The basic form of the game is given by the payoff matrix shown in Table 1 where both players have two decision alternatives: to cooperate or to defect. The rows show the strategies of player 1, and the columns indicate the strategies of player 2. For each corresponding strategy pair the first number in each position gives the payoff of player 1, and the second value is the payoff of player 2. The possible payoffs are R (reward for both agents cooperating), P (punishment for both agents defecting), T (temptation to defect when the other agent cooperates), and S (sucker payoff for cooperating when the other agent defects). The classic situation involves two suspects being arrested and separated by the police. There is insufficient evidence for the police to convict either one of the accused crime, so the police offers each one a deal to testify against the other. If both suspects cooperate with each other by remaining silent, each one is sentenced to minimal jail time on a lesser charge. If one suspect testifies against the other (defects) and the other stays silent (cooperates), then the defector is released and the cooperator is sentenced to full jail time on the accused charge. If both suspects defect and testify, then both are sentenced to reduced jail time on the accused charge. For the prisoner's dilemma the payoffs for the players satisfy the relation $T > R > P > S$ (Osborne 2004).

In this game defecting strictly dominates cooperating and thus the only equilibrium is for both players to defect. The dilemma is that both players would be better off if they both cooperated, but this is not a stable equilibrium since both players have a temptation to defect.

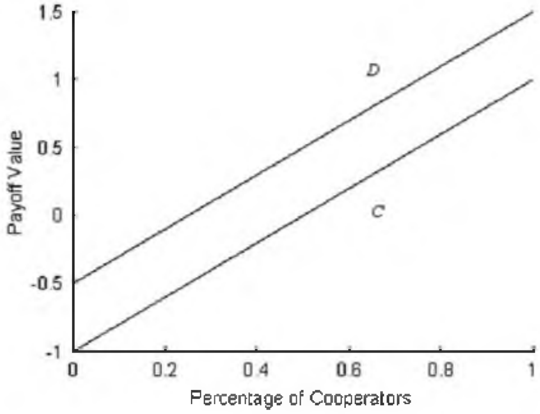
Study of collective behavior requires a multi-person extension of the model of this game.

In an N-person prisoner's dilemma game each agent can choose to cooperate or defect and then receives a reward or punishment that depends on the accumulated choices of the other agents. The amount of reward or punishment an agent receives is called a payoff function. A typical linear payoff function is shown in Fig. 1. In

Table 1 Payoff matrix for two player games

1	2	
	Cooperate	Defect
Cooperate	R, R	S, T
Defect	T, S	P, P

Fig. 1 Linear payoff functions for cooperators (*C*) and defectors (*D*)



this payoff function x is the percentage of cooperators, $C(x)$ is the payoff value for those agents that are cooperating and $D(x)$ is the payoff value for those agents that are defecting. The game is said to be an N-person prisoner's dilemma when the parameters of the payoff functions satisfy the relation $T > R > P > S$ (Schelling 1973) similarly to the two-person case.

A cartel would be an example of an N-person prisoner's dilemma game. In a cartel the players cooperate to set an optimum production and price as if it were a monopoly. Each player or agent agrees to charge the monopoly price and limit production to set forth quantities. The temptation for each agent is to defect by lowering the price and producing more. The defecting agent has a higher profit or payoff since its market share increases significantly while the remaining cooperating agents lose market share while holding the same monopoly price. If all agents defect, then all are worse off since they would all be charging the same lower price with the same market share as if they all would have cooperated.

Another interesting example of an N-person prisoner's dilemma game is given by the well known *tragedy of the commons*, (Hardin 1968). Consider a common ground used for the grazing of animals. When individuals add animals to their flocks they increase their personal wealth. Yet, every animal added to the total degrades the commons a small amount. The degradation for each additional animal is small relative to the gain in wealth for the owner, nevertheless if all owners follow this pattern the commons will ultimately be destroyed. When assuming rational actors, owners have incentive to increase their flock.

In an N-person game each agent decides to cooperate or defect depending on several factors; some authors have considered aspects such as reciprocal cooperation (Bornstein et al. 1994), others have considered fairness Fehr and Schmidt (1999). When abstracting from the reasons driving agents' decisions it is possible to categorize them as behavioral classes. In Szilagyi (2003) several behavioral classes are presented including Pavlovian, Greedy, Conformist and Accountant. This paper will concentrate on what in Szilagyi (2003) are referred as Pavlovian agents and other types will be only briefly discussed. This kind of agent has a certain probability of cooperating in each time period, which changes for the next time period or iteration by a proportion of the reward or punishment received;

$$p_i(t) = \begin{cases} p_i(t-1) + \alpha C(x(t-1)) & \text{if the agent cooperated at time period } t-1 \\ p_i(t-1) - \beta D(x(t-1)) & \text{if the agent defected at time period } t-1 \end{cases} \tag{1}$$

where $C(x(t-1))$ and $D(x(t-1))$ are respectively the payoffs for cooperators and defectors depending on the percentage $x(t-1)$ of cooperating agents in the population at time period $t-1$; α is the proportion or learning factor for cooperators and β is the learning factor for defectors.

With this updating rule the probability p of taking a certain action—C or D—changes by an amount proportional to its reward or penalty from the environment. In Szilagyi (2003) this update rule is referred as Pavlovian; yet, since the process in which learning occurs is a function of the consequences of behavior, it would be better to call to it *Skinnerian*. In fact, according to Skinner (1953, p. 65), "In operant conditioning we 'strengthen' an operant in the sense of making a response more probable or, in actual fact, more frequent". Operant conditioning framework was provided by the "law of effect" (Thordike 1911). The law of effect informs also other models of learning such as the implementation of Bush and Mosteller's stochastic learning provided in Macy and Flache (2002). By contrast, in the updating rule (1) we do not consider an aspiration level relative to which payoffs are positively or negatively evaluated on a standard scale, rather introduce different learning factors. In this paper, we refer to this class as Skinnerian even if the name under which it has been considered so far in the literature was Pavlovian (Szilagyi 2003). Finally, it must be noted that since this strategy uses learning factors it is not the same as the strategy Pavlov discussed in Nowak and Sigmund (1993).

The learning factors α and β are typically set to be the same value since in many models an agent is just responding to the reward or punishment, but α and β may be different in cases where the decision affects the information an agent receives for the next time period or includes other intangible qualities such as happiness of making a certain decision or even be the result of cultural factors. For example, a situation may occur in the cartel example when a cooperating agent receives more information from the other cooperating agents about the economical environment. In essence, defecting agents may be isolated and so may respond differently to a reward or punishment. It is an easier way to model this situation with different learning factors then by incorporating it into the payoff functions dealing with more quantifiable parameters such as production levels, market share, and profits. A similar case occurs when the agents are basically "happy" with being cooperators and are less likely to put forth a strong effort to reevaluate their probability after cooperating in any time period. They are simply following the flow. It could also be that defecting agents are "shunned" and react more strongly one way or another. These situations also can be modeled with different learning factors. As an extreme case let's assume that a defecting agent's decision is to isolate itself from the environment and enter the next iteration basically with the same state. For this extreme case the defecting agent would still receive a reward or punishment with zero learning factor where a cooperating (involved) agent may have some positive learning factor.

In Szilagyi (2003) it is shown that for Pavlovian agents an equilibrium occurs in an N-person prisoner's dilemma game when $x^* \, C(x^*) = (1 - x^*) \, D(x^*)$ under the assumption that the cooperating and defecting agents have the same learning factor. This equation depicts the situation when the total payoff of all the cooperating agents equals the total payoff of all the defecting agents. This equilibrium equation requires the solution of a quadratic equation with linear payoff functions.

In the next section we will develop a more general equilibrium equation that solves the N-person prisoner's dilemma game with different learning factors. The theoretical results will be also illustrated by simulation study.

3 Steady state analysis

An equation for the expected value of p will first be developed for an arbitrary size population of Skinnerian agents. To simplify the notation p_i^{new} denotes the probability that an agent i will be cooperating in the next time period, $C(x)$ is the payoff for cooperating agents and $D(x)$ is the payoff for the defecting agents. The probability that an agent i will be cooperating in the next period is given as

$$p_i^{\text{new}} = \begin{cases} p_i + \alpha C(x) & \text{if the agent is currently a cooperator} \\ p_i - \beta D(x) & \text{if the agent is currently a defector} \end{cases} \tag{2}$$

where x is the ratio of the agents who are currently cooperating in the population.

Since an agent is a cooperator with probability p_i and a defector with probability $1 - p_i$, the expected value of p_i^{new} becomes

$$\begin{aligned} \mathrm{E}(p_i^{\text{new}}) &= [p_i + \alpha C(x)]p_i + [p_i - \beta D(x)](1 - p_i) \\ &= \alpha p_i C(x) + p_i - \beta D(x) + \beta p_i D(x). \end{aligned} \tag{3}$$

If N is the total number of agents then the percentage of cooperating agents is $x = \frac{1}{N}\sum_{i=1}^{N} p_i$. By combining this relation and Eq. (3) we have

$$\begin{aligned} x^{\text{new}} &= \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}(p_i^{\text{new}}) \\ &= \sum_{i=1}^{N} \left[\frac{\alpha p_i C(x)}{N} + \frac{p_i}{N} - + \frac{\beta D(x)}{N} + \frac{\beta p_i D(x)}{N}\right] \\ &= [\alpha C(x) + \beta D(x) + 1]x - \beta D(x). \end{aligned} \tag{4}$$

At any steady state x^*, the state does not change anymore. That is, if $x = x^*$, then $x^{\text{new}} = x^*$ as well. Therefore from Eq. (4) we conclude that x^* is a steady state if and only if

$$x^* = [\alpha C(x^*) + \beta D(x^*) + 1]x^* - \beta D(x^*),$$

which can be written as

$$x^* \alpha C(x^*) = (1 - x^*) \beta D(x^*). \tag{5}$$

Note that if $\alpha = \beta$ then Eq. (5) simplifies to $x^*\, C(x^*) = (1 - x^*)\, D(x^*)$, which is the equilibrium equation presented in Szilagyi and Szilagyi (2002) for this special case.

The above analysis is valid in the case of general payoff functions since in the derivation no special forms are assumed. In the linear case $C(x) = S + (R - S)x$ and $D(x) = P + (T - P)x$ so Eq. (5) reduces to

$$x^* \alpha [S + (R - S)x^*] = (1 - x^*) \beta [P + (T - P)x^*].$$

This is a quadratic equation for x^* and therefore

$$x^* = \frac{-\alpha S - 2\beta P + \beta T \pm \sqrt{\alpha^2 S^2 - 2\alpha\beta ST + \beta^2 T^2 + 4\alpha\beta PR}}{2(\alpha R - \alpha S + \beta T - \beta P)}. \tag{6}$$

If $\alpha = \beta$, then this equation simplifies as

$$x^* = \frac{-S - 2P + T \pm \sqrt{S^2 - 2ST + T^2 + 4PR}}{2(R - S + T - P)}. \tag{7}$$

Equation (6) shows that the number of steady states is 0, 1 or 2 in the linear case depending on the signs of the discriminant and the roots which depend on the learning factors. On the contrary, in (7) the learning factors do not appear and the steady state depends only on the payoff values. In the latter case the identical learning factors have no influence on the theoretical prediction. If we assume that the learning factor is peculiar to the population culture, this would mean that different populations playing the same game would have the same steady states; this would be in contrast to Parks and Vu (1994) which showed differences in cooperation when comparing populations from different cultures. By contrast, when the learning factors are different their values determine the steady state of the population and the possible level of cooperation. This way, the learning factors characterize both the population and the steady states of the game they are playing. In this case the outcome of the game depends both on the game payoffs and the population playing it: this seems also to be more in accord to theories in which several entities are considered in the analysis of groups, see e.g. (Lewin 1947, p. 14).

Impulsive agents compare the common payoffs of cooperators $C(x)$ and that of defectors $D(x)$ where x is again the ratio of cooperators in the society. If $C(x) > D(x)$, then cooperators keep their previous strategy but defectors change it with a given probability of change δ_C. Otherwise the defectors keep their strategies and the cooperators become defectors with probability δ_D. In general $\delta_C \neq \delta_D$. Similarly to Eq. (4) it can be easily proved that the dynamic rule is the following:

$$x^{\text{new}} = \begin{cases} x + \delta_C(1 - x) & \text{if } C(x) > D(x) \\ x & \text{if } C(x) = D(x) \\ x - \delta_D x & \text{if } C(x) < D(x) \end{cases} \tag{8}$$

This generates a nonlinear discontinuous dynamics. Its properties are discussed in detail in Bischi et al. (2009a), and the relevant simulation study can be performed in the same way as shown in this paper. In Szilagyi (2003) impulsive agents are called greedy. Conformist and accountant agents have similar dynamic rules. The details are left as easy exercises to the reader.

4 Simulation study

We will confirm and extend the analysis of the previous section with different series of simulations. Firstly we will show that, the theoretical prediction can be obtained by simulation. Secondly, by introducing the bounds on probability values in order to keep them bounded we consider a nonlinear updating rule. Since the nonlinearity of this updating rule makes the analytical treatment of such dynamics impossible, simulation will allow us to extend these theoretical results to the nonlinear case. We performed the different simulations by a C++ program written by the authors, which is available from the authors upon request. For each case we report the average value obtained with 1,000 runs generated with different seeds for the random number generator. In particular, we followed the prescriptions reported in Press et al. (2007) and used random generator `Ran` (Press et al. 2007, p. 342). The results confirms those obtained with MATLAB and reported in Sandbank (2010).

In this paper interactions with all neighbors as a collective set are explored, i.e., the interaction is among the whole population. Time is moved forward in iterations. In each iteration the agents decide to cooperate or defect based on the probability distribution (1).

An example earlier introduced in Szilagyi (2003) will be reviewed. First, with the assumption that $\alpha = \beta$ the theoretical and simulation results are compared to those given in Szilagyi (2003). Next, the learning factor β is varied while keeping α the same so that we can examine cases where $\alpha \neq \beta$ and see the dependence of the steady state on the value of β.

The parameters given in Szilagyi (2003) are $\alpha = \beta = 0.1$, $P = -0.5$, $R = 1.0$, $S = -1.0$ and $T = 1.5$. We use the same parameters in order to compare our results to those in the literature.

Inserting α, β, P, R, S and T into Eq. (7) gives $x^* = 0.1798$ and 0.6952 which matches the results given in Szilagyi (2003). Figure 2 uses Eq. (4) to sequentially determine $x(t)$ from initial points of 0.05, 0.65 and 0.75. The quadratic curve is the right hand side of Eq. (4) and the straight line from (0,0) to (1,1) is a 45 degree line. For example, if we start at $x(0) = 0.05$ the next state is $x(1) = 0.0835$. This is represented by the vertical dashed line at $x(0) = 0.05$ running from 0 to 0.0835. Moving along the lowest horizontal line to $x(1) = 0.0835$ we come to the next vertical line which shows the second iteration going from $x(1) = 0.0835$ to $x(2) = 0.1071$. Continuing in this fashion we see that the solution converges to 0.1798. Figure 2 shows this procedure for starting states of 0.05, 0.65 and 0.75. From these results it is clear $x^* = 0.1798$ and 0.6952 are the solutions with $x^* = 0.1798$ being an attractor and $x^* = 0.6952$ being a repeller.

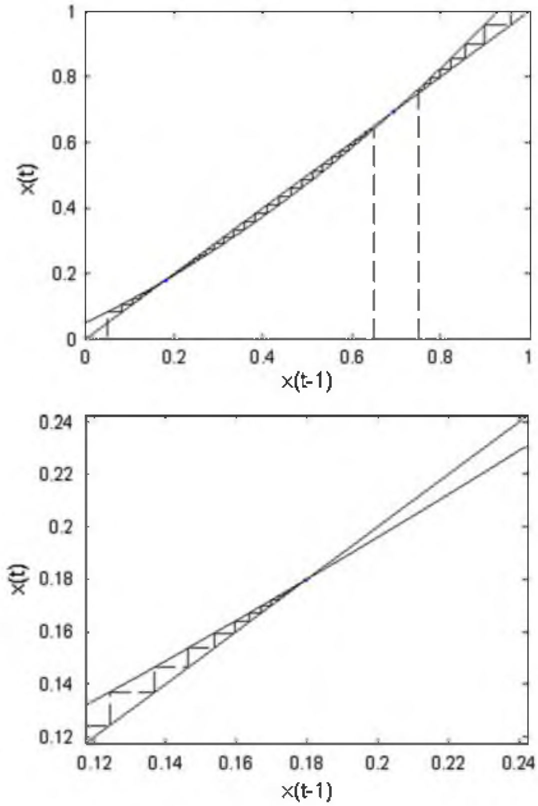


Fig. 2 The top graph is an iteration chart with initial starting states of 0.05, 0.65 and 0.75. This chart shows $x^* = 0.1798$ is an attractor and $x^* = 0.6952$ is a repeller. The chart on the bottom is a magnification using the initial starting state of 0.05. It better shows the iterations converging to the attractor $x^* = 0.1798$

In the simulation study we selected identical cooperation probabilities for all agents in the population. For each later time period we updated the cooperation probabilty of each agent by using Eq. (2). Based on these probability values the actual strategies (cooperation or defection) are generated. The percentage of the cooperative agents is the $x(t)$ value for this time period. This procedure is repeated for 500 time periods.

When using Eq. (1) to adjust the probability values, the simulation gave the same results as obtained by the analytical study. However there is a slight problem with this equation, since $p_i(t)$ might become infeasible as a probability value. So the equlibrium equation holds if all computed probability values are between zero and one. It is more realistic to make the following adjustment. If the computed $p_i(t)$ value becomes larger than one, then it is adjusted to one. Similarly, if its value becomes negative, then it is adjusted to zero. This adjustment process makes Eq. (1) nonlinear and nondifferentiable. Therefore, an equation similar to (4) cannot be obtained in this case. The dynamic properties of this more realistic model can be examined by using simulation.

Figure 3 shows some $x(t)$ sequences starting from different initial states. They coincide with the corresponding results presented in Szilagyi (2003). These results seem to contradict the theoretical findings. The presence of multiple solutions is the

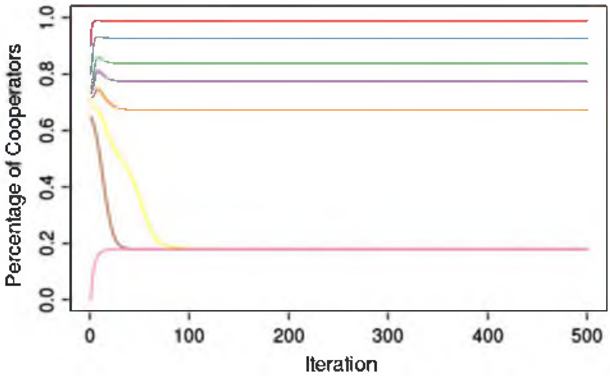


Fig. 3 500 iterations are performed considering the entire set of agents. $\alpha = \beta = 0.1$ are selected. The initial cooperating ratios from top to bottom curves are 0.90, 0.80, 0.75, 0.73, 0.71, 0.69, 0.65 and 0.00

result of adjusting too large probability values to one by decreasing them and similarly, adjusting negative probability values to zero by increasing them. In fact, when the initial cooperation ratio is below the repeller, the solution of the game converges toward the attractor where it stabilizes exactly. On the contrary, as mentioned in Szilagyi (2003), when the initial cooperation ratio is above the repeller the population does not result in the aggregate cooperation proportion converging to 1, as might be expected. In fact, when an individual agent starts off as a defector, it is possible that the agent will continue to defect as the reward for defecting drives its cooperation probability to zero.

We can vary the values of α and β and repeat the computation and simulation. In the following example we change the value of β to 0.01, so that $\alpha \neq \beta$. The parameter values in this case are $\alpha = 0.1$, $\beta = 0.01$, $P = -0.5$, $R = 1.0$, $S = -1.0$ and $T = 1.5$. Starting form the parameters used in Szilagyi (2003) we introduce a variation in the learning factor for defectors; the different learning factors for cooperators and defectors show why heterogeneity is important. We choose a smaller learning factor for defectors in order to consider a situation in which defectors tend to stick to their strategy. In future reasearch we will explore other learning parameters configurations.

There would be no change in inserting these parameters into Eq. (7) since this equation does not depend on α or β, but inserting them into the more general Eq. (6) leads to the different solutions of 0.0433 and 0.5249. The simulation results for $\alpha = 0.1$, $\beta = 0.01$ are shown in Figs. 4 and 5. Figure 4 uses the same initial probabilities as in Fig. 3 to show that the curves become different when $\alpha \neq \beta$. Figure 5 is based on the same data with different initial cooperating probabilities of 0.90, 0.80, 0.75, 0.56, 0.54, 0.52, 0.50 and 0 to show that the repeller is indeed 0.5249 as indicated by Eq. (6). It can also be seen from these simulations that the attractor is 0.0433, which is the same as determined by Eq. (6). This results show that Eq. (6) correctly determines the solution when $\alpha \neq \beta$.

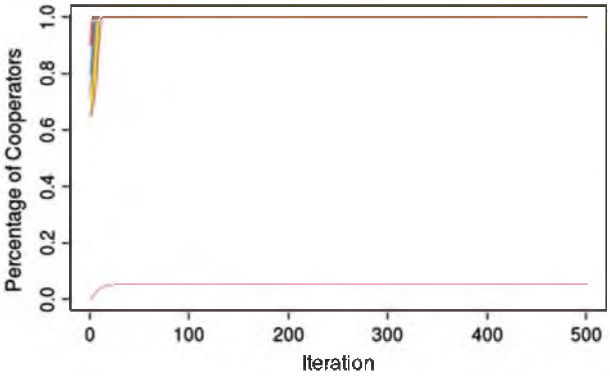


Fig. 4 500 iterations were performed considering the entire set of agents with $\alpha = 0.1$ and $\beta = 0.01$. The initial cooperating ratios from top to bottom curves are 0.90, 0.80, 0.75, 0.73, 0.71, 0.69, 0.65 and 0.00

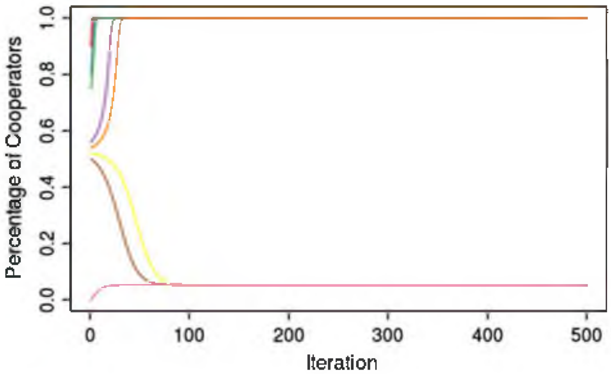


Fig. 5 This is the same simulation as Fig. 3 with $\alpha = 0.1$ and $\beta = 0.01$ and the initial cooperating ratios 0.90, 0.80, 0.75, 0.56, 0.54, 0.52, 0.50 and 0.00 from the top to the bottom. It can be seen here that the solutions of $x^* = 0.0433$ and 0.5249 agree with the analytical results of Eq. (6)

Table 2 Attractor solutions of Eq. (6) for varying learning factors

α	β			
	0.01	0.04	0.07	0.10
0.01	**0.1798**	0.2298	0.2383	0.2418
0.04	0.0892	**0.1798**	0.2064	0.2185
0.07	0.0584	0.1448	**0.1798**	0.1978
0.10	0.0433	0.1202	0.1582	**0.1798**

Bolded cells denote solutions for $\alpha = \beta$, which do not change

It is interesting to note that the solution does not change if both α and β vary but remains equal. Tables 2 and 3 shows the attractor and repeller solutions for the payoff function parameters used in the above examples as α and β vary. It is clear from these tables that the analytical solution does not change when the learning

Table 3 Repeller solutions of Eq. (6) for varying learning factors

α	β			
	0.01	0.04	0.07	0.10
0.01	**0.6953**	0.8702	0.9179	0.9401
0.04	0.5608	**0.6953**	0.7709	0.8172
0.07	0.5354	0.6280	**0.6953**	0.7433
0.10	0.5249	0.5940	0.6506	**0.6953**

Bolded cells denote solutions for $\alpha = \beta$, which do not change

factors are equal. This is understandable, since in Eq. (5) the multipliers α and β cancel out when $\alpha = \beta$.

5 Conclusion

An analytical equation, $x^*\alpha\ C(x^*) = (1 - x^*)\ \beta\ D(x^*)$, was found to solve the N-person prisoner's dilemma game with Skinnerian agents. This solution allows for situations where the agents have different learning factors and characterize the steady states of the dynamics depending on their learning factors which can be considered factors peculiar to the population. Through agent based simulation it was shown that this solution is accurate for linear payoff functions. If $\alpha = \beta$, then this analytical solution simplifies to $x^*\ C(x^*) = (1 - x^*)\ D(x^*)$ which is known from the literature. It was also shown by agent based simulation that the simplified equation $x^*\ C(x^*) = (1 - x^*)\ D(x^*)$ does not work for $\alpha \neq \beta$, which is also clear from the analytical solution. We briefly discussed the case of impulsive agents. We also discussed the limitations of the analytic equations and introduced a modified model eliminating those restrictions. The dynamics of this more realistic model was examined by simulation. In this paper some major simplifying assumptions were made. As the approach of integrating experimental data into theoretical and simulation models is becoming important in the recent literature, see for example (Dal Forno and Merlone 2004, 2012; Ebenhöh 2006; Boero et al. 2010), it will be interesting to analyze how the assumptions we made are realistic. Furthermore, relaxing these assumptions after observing human participants interact in situations similar to those here described will increase the predictive accuracy of the model. Also, in our future research we plan to address cases of nonlinear payoff functions with one or more intersection as those described in Schelling (1973). Also it will be interesting to study local effects, by considering more general neighborhood structures of the agents: when neighborhoods, in which the interaction takes place, are special subsets of the entire society. Then it will be possible to analyze local effects instead of global properties. Other agent types were not examined in detail in the literature yet. We also plan to investigate such cases in order to compare the theoretical findings to the simulation study results. When the comparison is possible it will help to obtain *vertical multiple implementations* of the model, i.e., using different modeling paradigms to understand a phenomena as suggested in Merlone et al. (2008). On the other hand, simulation will help to extend the results when heterogeneity of agents makes the formal derivation of results unfeasible.

References

Axelrod R (1984) The evolution of cooperation. Basic Books, New York

Bischi GI, Gardini L, Merlone U (2009a) Impulsivity in binary choices and the emergence of periodicity. Discret Dyn Nat Soc 407913, pp 22. doi:10.1155/2009/407913

Bischi GI, Gardini L, Merlone U (2009b) Periodic cycles and bifurcation curves for one-dimensional maps with two discontinuities. J Dyn Syst Geom Theor 7(2):101–123

Bischi GI, Merlone U (2009) Global dynamics in binary choice models with social influence. J Math Sociol 33(4):277–302

Boero R, Bravo G, Castellani M, Squazzoni F (2010) Why bother with what others tell you? An experimental data-driven agent-based model. J Artif Soc Soc Simul 13(3):6. ISSN 1460-7425. http://jasss.soc.surrey.ac.uk/13/3/6.html

Bornstein G, Erev I, Goren H (1994) The effect of repeated play in the IPG and IPD team games. J Confl Resolut 38(4):690–707

Brams S, Kilgour DM (1988) Game theory and national security. Blackwell, New York

Conybeare JA (1984) Public goods, prisoners' dilemmas and the international political economy. Int Stud Quart 28(1):5–22

Dal Forno A, Gardini L, Merlone U (2012) Ternary choices in repeated games and border collision bifurcations. Chaos Solitons Fractals 45(3):294–305. doi:10.1016/j.chaos.2011.12.003

Dal Forno A, Merlone U (2004) From classroom experiments to computer code. J Artif Soc Soc Simul 7(3):2. ISSN 1460-7425. http://jasss.soc.surrey.ac.uk/7/3/2.htlm

Dal Forno A, Merlone U (2012) Grounded theory based agents. In: Laroque C, Himmelspach J, Pasupathy R, Rose O, Uhrmacher AM (eds) Proceedings of the 2012 winter simulation conference

Dal Forno A, Merlone U (2013) Collision bifurcations in a model of Braess paradox. Math Comput Simul. doi:10.1016/j.matcom.2012.12.001

Dawes RM (1980) Social dilemmas. Annu Rev Psychol 31:169–193

Dawes RM, Messick DM (2000) Social dilemmas. Int J Psychol 35(2):111–116

Ebenhöh E (2006) Modeling non-linear common-pool resource experiments with boundedly rational agents. In: Sichman JS, Antunes L (eds) MABS 2005, number 3891 in LNAI. Springer, Berlin, pp 133–146

Epstein JM, Axtell R (1996) Growing artificial societies: social science from the bottom up. MIT Press, Cambridge

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. Q J Econ 114(3):817–868

Gardini L, Merlone U, Tramontana F (2011) Inertia in binary choices: continuity breaking and big-bang bifurcation points. J Econ Behav Organ 80:153–167

Gilbert N, Terna P (2000) How to build and use agent-based models in social science. Mind Soc 1:57–72

Hardin G (1968) The tragedy of commons. Science 162:1243–1248

Hauert C, Szabo G (2003) Prisoner's dilemma and public goods games in different geometries: compulsory versus voluntary interactions. Complexity 8:31–38

Kollock P (1998) Social dilemmas: the anatomy of cooperation. Annu Rev Sociol 24:183–214

Ledyard JO (1995) Public goods: a survey of experimental research. In: Kagel J, Roth A (eds) Handbook of experimental economics. Princeton University Press, Princeton, pp 89–108

Lewin K (1947) Frontiers in group dynamics. Concept, method and reality in social science; social equilibria and social change. Hum Relat I:5–41

Macy MW, Flache A (2002) Learning dynamics in social dilemmas. Proc Natl Acad Sci USA 99:7229–7236

Merlone U, Sonnessa M, Terna P (2008) Horizontal and vertical multiple implementations in a model of industrial districts. J Artif Soc Soc Simul 11(2):5. ISSN 1460-7425. http://jasss.soc.surrey.ac.uk/11/2/5.htlm

Merlone U, Szidarovszky F, Szilagyi MN (2007) Finite neighborhood games with binary choices. Mathematica Pannonica 18(2):205–217

Nowak M, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature 363:56–58
Osborne MJ (2004) An introduction to game theory. Oxford University Press, New York
Ostrom E (1998) A behavioral approach to the rational choice theory of collective action: presidential address, American Political Science Association, 1997. Am Polit Sci Rev 92(1):1–22
Ostrom E (2000) Collective action and the evolution of social norms. J Econ Perspect 14(3):137–158
Parks CD, Vu AD (1994) Social dilemma behavior of individuals from highly individualist and collectivist cultures. J Confl Resolut 38(4):708–718
Press WH, Teulkolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes in C++, 3rd edn. Cambridge University Press, Cambridge
Sandbank D (2010) Analytical solution, agent behavioral transitions and classification structures in *n*-person social dilemma games. PhD thesis, Systems and Industrial Engineering Department. The University of Arizona, Tucson, AZ
Santos FC, Santos MD, Pacheco JM (2008) Social diversity promotes the emergence of cooperation in public goods games. Nature 454:213–216
Schelling TC (1973) Hockey helmets, concealed weapons, and daylight saving. J Confl Resolut 17:381–428
Skinner BF (1953) Science and human behavior. MacMillan, New York
Szilagyi MN (2003) An investigation of *N*-person prisoners' dilemmas. Complex Syst 14(2):155–174
Szilagyi MN, Szilagyi ZC (2002) Non-trivial solutions to the *N*-person prisoners' dilemma. Syst Res Behav Sci 19(3):281–290
Tabarrok A (1998) The private provision of public goods via dominant assurance contracts. Public Choice 96(3–4):345–62
Thordike EL (1911) Animal intelligence. Hafner, Darien
Young J (1998) Using computer models to study the complexities of human society. Chron High Educ Sect A 4(46):17–19