



Giulio Cimini and Angel Sanchez (2015)

How Evolutionary Dynamics Affects Network Reciprocity in Prisoner's Dilemma

Journal of Artificial Societies and Social Simulation 18 (2) 22

<<http://jasss.soc.surrey.ac.uk/18/2/22.html>>

Received: 20-Jun-2014 Accepted: 06-Nov-2014 Published: 31-Mar-2015

Abstract

Cooperation lies at the foundations of human societies, yet why people cooperate remains a conundrum. The issue, known as network reciprocity, of whether population structure can foster cooperative behavior in social dilemmas has been addressed by many, but theoretical studies have yielded contradictory results so far—as the problem is very sensitive to how players adapt their strategy. However, recent experiments with the prisoner's dilemma game played on different networks and in a specific range of payoffs suggest that humans, at least for those experimental setups, do not consider neighbors' payoffs when making their decisions, and that the network structure does not influence the final outcome. In this work we carry out an extensive analysis of different evolutionary dynamics, taking into account most of the alternatives that have been proposed so far to implement players' strategy updating process. In this manner we show that the absence of network reciprocity is a general feature of the dynamics (among those we consider) that do not take neighbors' payoffs into account. Our results, together with experimental evidence, hint at how to properly model real people's behavior.

Keywords:

Evolutionary Game Theory, Prisoner's Dilemma, Network Reciprocity

Introduction

- 1.1 Cooperation and defection represent the two alternative choices behind social dilemmas (Dawes 1980). Cooperative individuals contribute to the well being of the community at their own expenses, whereas, defectors neglect doing so. Because of that cost of contribution, cooperators get lower individual fitness and thus selection favors defectors. This situation makes the emergence of cooperation a difficult matter. Evolutionary game theory (Maynard Smith & Price 1973) represents a theoretical framework suitable to tackle the issue of cooperation among selfish and unrelated individuals. Within this framework, social dilemmas are formalized at the most basic level as two-person games, where each player can either choose to cooperate (C) or to defect (D). The *Prisoner's Dilemma* game (PD) (Axelrod 1984) embodies the archetypal situation in which mutual cooperation is the best outcome for both players, but the highest individual benefit is given by defecting. Mathematically, this is described by a matrix of payoffs (entries correspond to the row player's payoffs) of the form:

	C	D
C	R	S
D	T	P

so that mutual cooperation bears R (reward), mutual defection P (punishment), and with the mixed choice the cooperator gets S (sucker's payoff) and the defector T (temptation). The heart of the dilemma resides in the condition $T > R > P > S$: both players prefer the opponent to cooperate, but the temptation to cheat ($T > R$) and the fear of being cheated ($S < P$) pull towards choosing defection: according to Darwinian selection, cooperation extinction is then unavoidable (Hofbauer & Sigmund 1998)—a scenario known as the *tragedy of the commons* (Hardin 1968).

- 1.2 However, cooperation is widely observed in biological and social systems (Maynard Smith & Szathmari 1995). The evolutionary

origin of such behavior hence remains a key unsolved puzzle across several disciplines, ranging from biology to economics. Different mechanisms have been proposed as putative explanations of the emergence of cooperation (Nowak 2006), including the existence of a social or spatial structure that determines the interactions among individuals—a feature known as *network reciprocity*. In a pioneering work, Nowak and May (1992) showed that the behavior observed in a repeated PD was dramatically different on a lattice than in a well-mixed population (or, in more physical terms, in a mean-field approach): in the first case, cooperators were able to prevail by forming clusters and preventing exploitation from defectors. Subsequently, many researchers devoted their attention to the problem of cooperation on complex networks, identifying many differences between structured and well-mixed populations (Roca et al. 2009a) that by no means were always in favor of cooperation (Hauert & Doebeli 2004; Sysi-Aho et al. 2005; Roca et al. 2009b). The main conclusion of all these works is that this problem is very sensitive to the details of the system (Szabó & Toke 2007; Roca et al. 2009a, 2009b), in particular to its evolutionary dynamics (Hofbauer & Sigmund 1998, 2003), i.e., the manner in which players adapt their strategy. On the experimental side, tests of the different models were lacking (Helbing & Yu 2010), because the few available studies (Cassar 2007; Kirchkamp & Nagel 2007; Traulsen et al. 2010; Fischbacher et al. 2001; Suri & Watts 2011) dealt only with very small networks. Network sizes such that clusters of cooperators could form have been considered only in recent large-scale experiments (Grujić et al. 2010; Gracia-Lázaro et al. 2012b) with humans playing an iterated multiplayer PD, as in the theoretical models. The outcome of the experiments was that, when it comes to human behavior, the existence of an underlying network of contacts does not have any influence on the global outcome.

- 1.3 The key observation to explain the discrepancy between theory and experiments is that most of the previous theoretical studies have been building on evolutionary dynamics based on payoff comparison (Hofbauer & Sigmund 2003; Roca et al. 2009a). While these rules are appropriate to model biological evolution (with the payoff representing fitness and thus reproductive success), they may not apply to social or economic contexts—where individuals are aware of others' actions but often do not know how much they benefit from them. Also when the latter information is available, recent analysis (Grujić et al. 2014) of experimental outcomes (Fischbacher et al. 2001; Traulsen et al. 2010; Grujić 2010; Gracia-Lázaro 2012b) show that humans playing PD or Public Good games do not base their decisions on others' payoffs. Rather, they tend to reciprocate the cooperation that they observe, being more inclined to contribute the more their partners do. The independence on the topology revealed in Gracia-Lázaro et al. (2012b) can be therefore seen as a consequence of this kind of behavior (Gracia-Lázaro et al. 2012a). Notably, absence of network reciprocity has also been observed in theoretical studies based on Best Response dynamics (an update rule that, as we will see, is independent on neighbors' payoffs) (Roca et al. 2009c) and in a learning-based explanation of observed behaviors (Cimini and Sánchez 2014). This suggests that the absence of network reciprocity in the iterated PD may be general for any evolutionary dynamics that does not take neighbors' payoffs into account.
- 1.4 In this paper we aim specifically at shedding light on this point. In order to do so, we develop and study an agent-based model of a population of individuals, placed on the nodes of a network, who play an iterated PD game with their neighbors—a setting equivalent to that of recent experiments (Traulsen et al. 2010; Grujić et al. 2010; Gracia-Lázaro et al. 2012b)—and whose strategies are subject to an evolutionary process. The key point in this work is that we consider a large set of evolutionary dynamics, representing most of the alternatives that have been proposed so far to implement the strategy updating process. At the same time, we consider a large set of population structures, covering most of the studied models of complex networks. In this way we are able to make, on the same system, a quantitative comparison of the different evolutionary dynamics, and check the presence of network reciprocity in the different situations. At the end we show that the absence of network reciprocity is a general consequence of evolutionary dynamics that are not based on payoff comparison.
- 1.5 Note that our study here is intrinsically different from the work of Cimini and Sánchez (2014), where we have investigated the issue of whether the behavior observed in the experiments (Grujić et al. 2010; Gracia-Lázaro 2012b)—named *moody conditional cooperation* or MCC—could arise from an evolutionary dynamics acting on the MCC behavioral parameters themselves, and if the resulting evolution of the system was compatible with the experimental outcomes. The work reported in Cimini and Sánchez (2014) was therefore only intended to study whether MCC might have an evolutionary explanation. On the contrary, as explained below, here we are considering stochastic behaviors ruled by mixed strategies (i.e., probabilities of cooperating) that are subject to an evolutionary process—and not individuals using MCC-like strategies. This approach is indeed the traditional one encountered in the literature; with this choice, while we do not aim at reproducing the experimental results of Grujić et al. (2010) and Gracia-Lázaro et al. (2012b), we are able to make an extensive comparison of the evolutionary dynamics that are most commonly used, and detect the effects of the population structure and the presence of network reciprocity in the different dynamics to which mixed strategies might be subjected.



Model

- 2.1 We consider a population of N individuals, placed on the nodes of a network and playing an iterated PD game with their neighbors. During each round t of the game, each player chooses to undertake a certain action (C or D) according to her strategy profile, then plays a PD game with her k neighbors (the selected action remains the same with all of them) and finally receives the corresponding payoff π^t —which is the sum of the single payoffs obtained in all her k pairwise games. Each player's strategy is modeled stochastically by the probability $p^t \in [0, 1]$ of cooperating at round t . Strategies are subject to an evolutionary process, meaning that every τ rounds each player may update her probability of cooperating according to a particular rule. Note that each player starts with an initial probability of cooperating $p^{(0)}$ drawn from a uniform distribution $Q[p^{(0)}] = \mathcal{U}[0, 1]$. This results in an

expected initial fraction of cooperators c_0 equal to $1/2$ —a value close to what is observed in the experiments (Grujić et al. 2010; Gracia-Lázaro et al. 2012b), and otherwise representing our ignorance about the initial strategy of the players. In any event, our results remain valid for any (reasonable) form of the distribution $Q[\rho^{(0)}]$.

2.2 We consider different parameterizations for the PD game, i.e., different intensities of the social dilemma. While we leave $R=1$ and $P=0$ fixed, we take T values in the range $(1,2)$, and S values in the range $(-1,0]$ (note that $S=P=0$ corresponds to the "weak" PD). More importantly, we take into account different patterns of interactions among the players. These include the "well-mixed" population, represented by an Erdős–Rényi random graph with average degree m , rewired after each round of the game (which we indicate as *well-mixed*), as well as static networks (all with average degree m): Erdős–Rényi random graphs (*random*), scale free random networks with degree distribution $P(k) \approx k^{-3}$ (*scale-free*) and regular lattice with periodic boundary conditions—where each node is linked to its $k=m$ nearest neighbors (*lattice*). Finally, we include two real instances of networks, the first given by the e-mail interchanges between members of the University Rovira i Virgili in Tarragona (*email*) (Guimerà et al. 2003), and the second being the giant component of the user network of the Pretty-Good-Privacy algorithm for secure information interchange (*PGP*) (Boguñá et al. 2004). The degree distributions of all these networks are reported in Figure 1. In simulations, we build the artificial networks using $N=1000$ and $m=10$,^[1] whereas, for the two real networks it is $N=1133$, $m=19.24$ (*email*) and $N=10679$, $m=4.56$ (*PGP*).

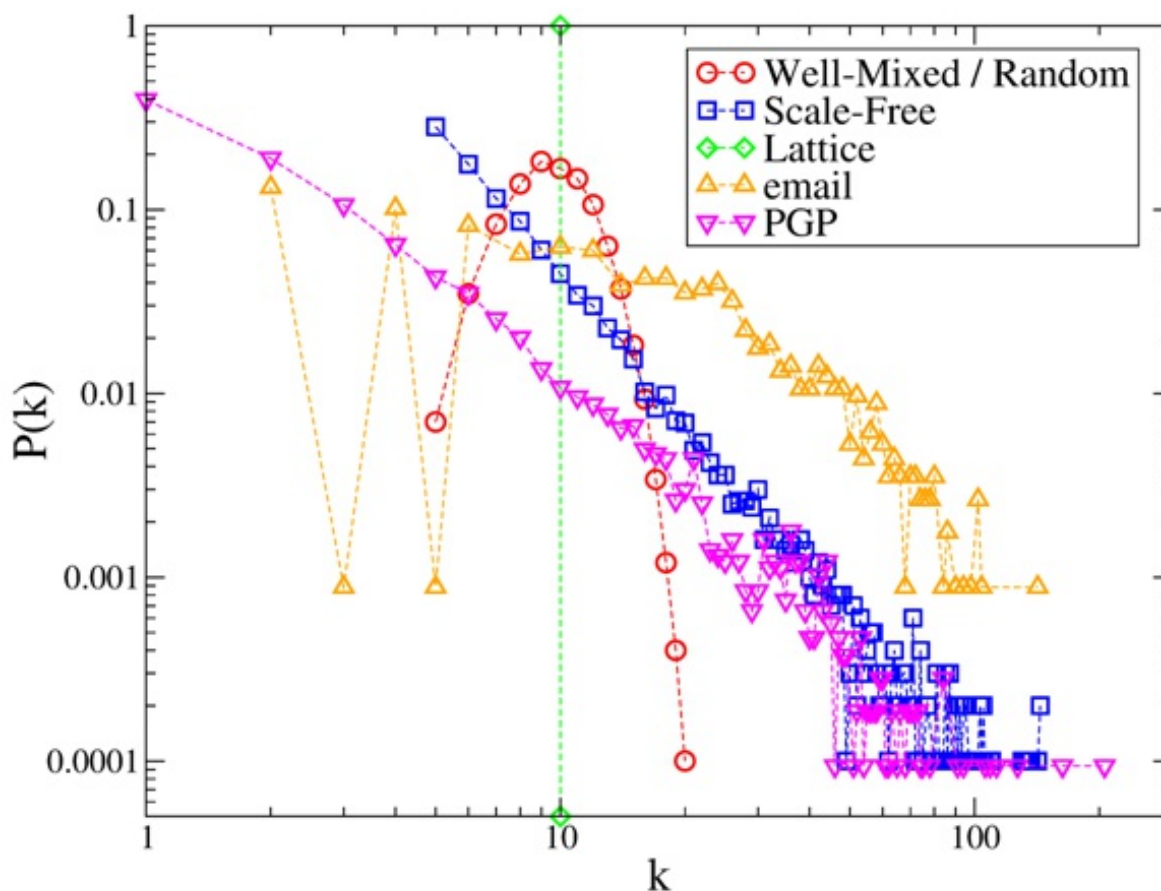


Figure 1. Degree distributions $P(k)$ of the considered networks. Note that for the well-mixed case the network is dynamic but $P(k)$ does not change, and thus is identical to that of random static networks.

2.3 The original and most important aspect of this study is that we consider a large variety of evolutionary dynamics for players to update their strategies, covering the rules that are most frequently employed in the literature to model the strategy updating process. Note that we do not consider *reactive strategies* (Sigmund 2010)—i.e., probabilities to cooperate that depend directly on the opponents' previous actions; rather, we focus on the simplest case of modeling strategies as unconditional propensities to cooperate (which, however, contain past actions through the evolutionary dynamics). Yet, our study is more general than most of those encountered in the literature—where only pure strategies (i.e., playing always C or D) are considered. In our framework of mixed strategies, pure strategies can arise as special (limit) cases, when for a player p becomes equal to 0 or 1.

2.4 We start by exploring a set of rules of imitative nature, representing a situation in which bounded rationality or lack of information force players to copy (imitate) others' strategies (Schlag 1998). These rules are widely employed in the literature to model evolutionary dynamics. Here we consider the most notorious ones, in which a given player i may adopt a new strategy by copying the probability of cooperating p from another player j , which is one of the k_i neighbors of i .

2.5 *Proportional Imitation* (Helbing 1992)— j is chosen at random, and the probability that i imitates j depends on the difference

between the payoffs that they obtained in the previous iteration of the game:

$$\begin{aligned} \text{Prob}\{\rho_j^{(t)} \rightarrow \rho_j^{(t+1)}\} &= (\pi_j^{(t)} - \pi_i^{(t)}) / \Phi_{ij} \text{ if } \pi_j^{(t)} > \pi_i^{(t)} \text{ and} \\ \text{Prob}\{\rho_j^{(t)} \rightarrow \rho_j^{(t+1)}\} &= 0 \text{ otherwise.} \end{aligned}$$

Here, $\Phi_{ij} = \max(k_i, k_j) [\max(R, T) - \min(P, S)]$ to have $\text{Prob}\{\cdot\} \in [0, 1]$. This updating rule is known to bring (for a large, well-mixed population) to the evolutionary equation of the replicator dynamics (Schlag 1998).

- 2.6 *Fermi rule* (Szabó & Toke 1998)— j is again chosen at random, but the imitation probability depends now on the payoff difference according to the Fermi distribution function:

$$\text{Prob}\{\rho_j^{(t)} \rightarrow \rho_j^{(t+1)}\} = 1 / \{1 + \exp[-\beta(\pi_j^{(t)} - \pi_i^{(t)})]\}.$$

Note that mistakes are possible under this rule: players can imitate others who are gaining less. The parameter β regulates selection intensity, and is equivalent to the inverse of noise in the update rule.

- 2.7 *Death-Birth rule*, inspired by Moran dynamics (Moran 1962)—player i imitates one of her neighbors j , or herself, with a probability proportional to the payoffs:

$$\text{Prob}\{\rho_j^{(t)} \rightarrow \rho_j^{(t+1)}\} = (\pi_j^{(t)} - \psi) / [\sum_{k \in n_i^*} (\pi_k^{(t)} - \psi)],$$

where n_i^* is the set including i and her neighbors and $\psi = \max_{j \in n_i^*} k_j \min(0, S)$ to have $\text{Prob}\{\cdot\} \in [0, 1]$. As with the Fermi rule, mistakes are allowed here.

- 2.8 *Unconditional Imitation* or "Imitate the Best" (Nowak 1992)^[2]—under this rule each player i imitates the neighbor j with the largest payoff, provided this payoff is larger than the player's:

$$\text{Prob}\{\rho_j^{(t)} \rightarrow \rho_j^{(t+1)}\} = 1 \quad \text{if } j: \pi_j^{(t)} = \max_{k \in n_i^*} \pi_k^{(t)}.$$

Note that while the first three rules are stochastic, Unconditional Imitation leads to a deterministic dynamics.

- 2.9 *Voter model* (Holley & Liggett 1975)— i simply imitates a randomly selected neighbor j . Differently from the other imitative dynamics presented here (in which the imitation mechanism is based on the payoffs obtained in the previous round of the game), the Voter model is not based on payoff comparison, but rather on social pressure: players simply follow the social context without any strategic consideration (Fehr & Gächter 2000; Vilone et al. 2012, 2014).

- 2.10 We also consider two evolutionary dynamics that go beyond pure imitation and are innovative, allowing extinct strategies to be reintroduced in the population (whereas imitative dynamics cannot do that). As we will see, neither of these rules (nor the Voter model) makes use of the information on others' payoffs.

- 2.11 *Best Response* (Matsui 1992; Blume 1993)—this rule has been widely employed in economic contexts, embodying a situation in which players are rational enough to compute the optimum strategy, i.e., the "best response" to what others did in the last round. More formally, at the end of each round t a given player i uses $x_i^{(t)}$ (the fraction of neighbors who cooperated at t) to compute the payoffs that she would have obtained by having chosen action C or D, respectively:

$$E[\pi_i^{(t)}(C)] = R x_i^{(t)} + S(1 - x_i^{(t)}); \quad E[\pi_i^{(t)}(D)] = T x_i^{(t)} + P(1 - x_i^{(t)}).$$

The quantity to increase is then:

$$E[\pi_i^{(t)}] = \rho_i^{(t)} E[\pi_i^{(t)}(C)] + (1 - \rho_i^{(t)}) E[\pi_i^{(t)}(D)].$$

The new strategy $\rho_i^{(t+1)}$ is picked among $\{\rho_i^{(t)}, \rho_i^{(t)} + \delta, \rho_i^{(t)} - \delta\}$ (where δ is the "shift") as the one that brings to the highest $E[\pi_i^{(t)}]$ (and satisfies $0 \leq \rho_i^{(t+1)} \leq 1$). Note that we do not use exhaustive best response here (which consists in choosing $\rho_i^{(t+1)}$ as the value of p that maximizes $E[\pi_i^{(t)}]$) as it leads immediately to the Nash equilibrium of the PD game ($\rho_i = 0 \forall i$). Best Response belongs to a family of strategy updating rules known as *Belief Learning* models, in which players update beliefs about what others will do according on accumulated past actions, and then use those beliefs to determine the optimum strategy. Best Response is a limit case that uses only last round actions to determine such optimum. We chose to restrict our attention to Best Response for three main reasons. 1) it allows for a fair comparison with the other updating rules, that only rely on last round information; 2) in the non-exhaustive formulation of Best Response, history is held in the actual values of the parameter p ; 3) in the one-shot, strong PD the Nash equilibrium is full defection, hence at the end the system collapses to this state for any information used to build beliefs about others' actions.

- 2.12 *Reinforcement Learning* (Bush & Mosteller 1955; Macy & Flache 2002; Izquierdo et al. 2008; Cimini & Sánchez 2014)—under this rule, a player uses her experience to choose or avoid certain actions based on their consequences: choices that met

aspirations in the past tend to be repeated in the future, whereas, choices that led to unsatisfactory outcomes tend to be avoided. This dynamics works as follows. First, after each round t , player i calculates her "stimulus" $s_i^{(t)}$ as:

$$s_i^{(t)} = (\pi_i^{(t)} / k_i - A_i^{(t)}) / \max\{|T - A_i^{(t)}|, |R - A_i^{(t)}|, |P - A_i^{(t)}|, |S - A_i^{(t)}|\},$$

where $A_i^{(t)}$ is the current "aspiration level" of player i , and normalization assures $|s_i^{(t)}| \leq 1$. Second, each player updates her strategy as:

$$p_i^{(t+1)} = p_i^{(t)} + \lambda s_i^{(t)} (1 - p_i^{(t)}) \text{ if } s_i^{(t)} > 0 \text{ and } p_i^{(t+1)} = p_i^{(t)} + \lambda s_i^{(t)} p_i^{(t)} \text{ otherwise.}$$

Here, $\lambda \in (0, 1]$ is the learning rate—low and high λ s representing slow and fast learning, respectively (hence for simplicity we use $\tau = 1$ in this case). Finally, player i can adapt her aspiration level as:

$$A_i^{(t+1)} = (1 - h) A_i^{(t)} + h \pi_i^{(t)} / k_i$$

where $h \in [0, 1)$ is the adaptation (or habituation) rate. Note that, when learning, players rely only on the information about their own past actions and payoffs.

2.13 In summary, our model is defined as follows:

- We place N players on the nodes of the chosen social network; for each player, we draw the initial probability of cooperating $p^{(0)}$ from the initial distribution $Q[p^{(0)}] = \mathcal{U}[0, 1]$.
- Each step t of the simulation is a round of the game, in which each player i chooses action C with probability $p_i^{(t)}$ (and action D otherwise), then plays a one-shot PD game with her k_i neighbors and collects the resulting payoff $\pi_i^{(t)}$ —defined as the sum of the single payoffs obtained in all her k_i pairwise games. These actions happen simultaneously for all players.
- Every τ rounds players simultaneously update their probability of cooperating according to one of the evolutionary rules described above.

The pseudo-code for the model is available at <https://www.openabm.org/model/4354/version/1/view>.



Results and Discussion

3.1 Here we report the results of the extensive simulation program for the model described above. In the following discussion, we focus our attention on the particular evolutionary dynamics employed, as well as on the specific network topology describing the interaction patterns among the players. We study the evolution of the level of cooperation c (i.e., the fraction of cooperative players in each round of the game), as well as the stationary probability distribution $Q^*(p)$ of the individual strategies (i.e., the parameters $\{p\}$) among the population. We will show results relative to the case $\tau = 10$ (we update players' strategy every ten rounds), yet we have observed that the particular value of τ influences the convergence time of the system to its stationary state, but does not alter its qualitatively characteristics. Also, we will report examples for two sets of game parameters ($T = 3/2$, $S = -1/2$ and $T = 3/2$, $S = 0$) but our findings are valid for the whole range studied—the main differences appearing between the "strong" and "weak" version of the PD game. Finally, we will show results averaged over a low number of realizations because, as the experiments show, the absence of network reciprocity is observed for single realizations, and hence it should be recovered from the model in the same manner.

Imitative dynamics

- 3.2 Results for the different imitation-based strategy updating rules are reported in Figures 2 and 3. As plots clearly show, the final level of cooperation in this case depends heavily on the population structure, and often the final state is full defection (especially for the well-mixed case). The fact that these updating are imitative and not innovative generally leads, for an individual realization of the system, to a very low number of strategies p (often just one) surviving at the end of the evolution. However, the surviving strategies are indeed different among independent realizations (but when $p \rightarrow 0$, i.e., the final outcome is a fully defective state). This points out to the absence of strategies that are stable attractors of the evolution.
- 3.3 The easiest situation to understand is perhaps given by employing the Voter model as the update rule: since there is no mechanism here to increase the payoffs, the surviving strategy is just randomly selected among those initially born in the population, and thus the average cooperation level is determined by $Q[p^{(0)}]$ (the probability distribution of the initial parameters p). This happens irrespectively of the particular values of T and S and, more importantly, of the specific topology of the underlying social network. A similar situation is observed with the Fermi rule for low β (high noise). Indeed, in this case errors are frequent, so that players copy the strategies of others at random and c remains close on average to its initial value c_0 . The opposite limit of high β (small noise, the case reported in the plots) corresponds instead to errors occurring rarely, meaning that players always

copy the strategy of others who have higher payoffs. In the majority of cases, for the strong PD this leads to a fully defective final outcome. The exception is given by games played on network topologies with broad degree distribution, where cooperation may thrive at a local scale (resulting in a small, non-zero value of c) because of the presence of hubs—see below for a more detailed discussion of this phenomenon. On the other hand, the weak PD showcases more diverse final outcomes: the stationary value of c is higher than c_0 for scale-free topologies, and a non-zero level of cooperation arises also in static random graphs and lattices. Note that in general we observe that the stationary (non-zero) cooperation levels decrease/increase for increasing/decreasing values of the temptation T , however such variations do not alter qualitatively the picture we present here—for this reason, we only present results for $T=3/2$. Moving further, Proportional Imitation leads to final outcomes very similar to those of the Fermi Rule for high β , which makes sense as the two rules are very similar—the only difference being the form of the updating probability (linear in the payoff difference for Proportional Imitation, highly non-linear for the Fermi rule). The Death-Birth rule and Unconditional Imitation also bring to similar results, and this is also due to their similarity in preferentially selecting the neighbor with the highest payoff. For these latter two rules, cooperation emerges for games played on all kinds of static networks (i.e., it does not only for a well-mixed population), with a stationary value of c that varies depending on the specific network topology and on the particular entries of the payoff matrix (T and S).

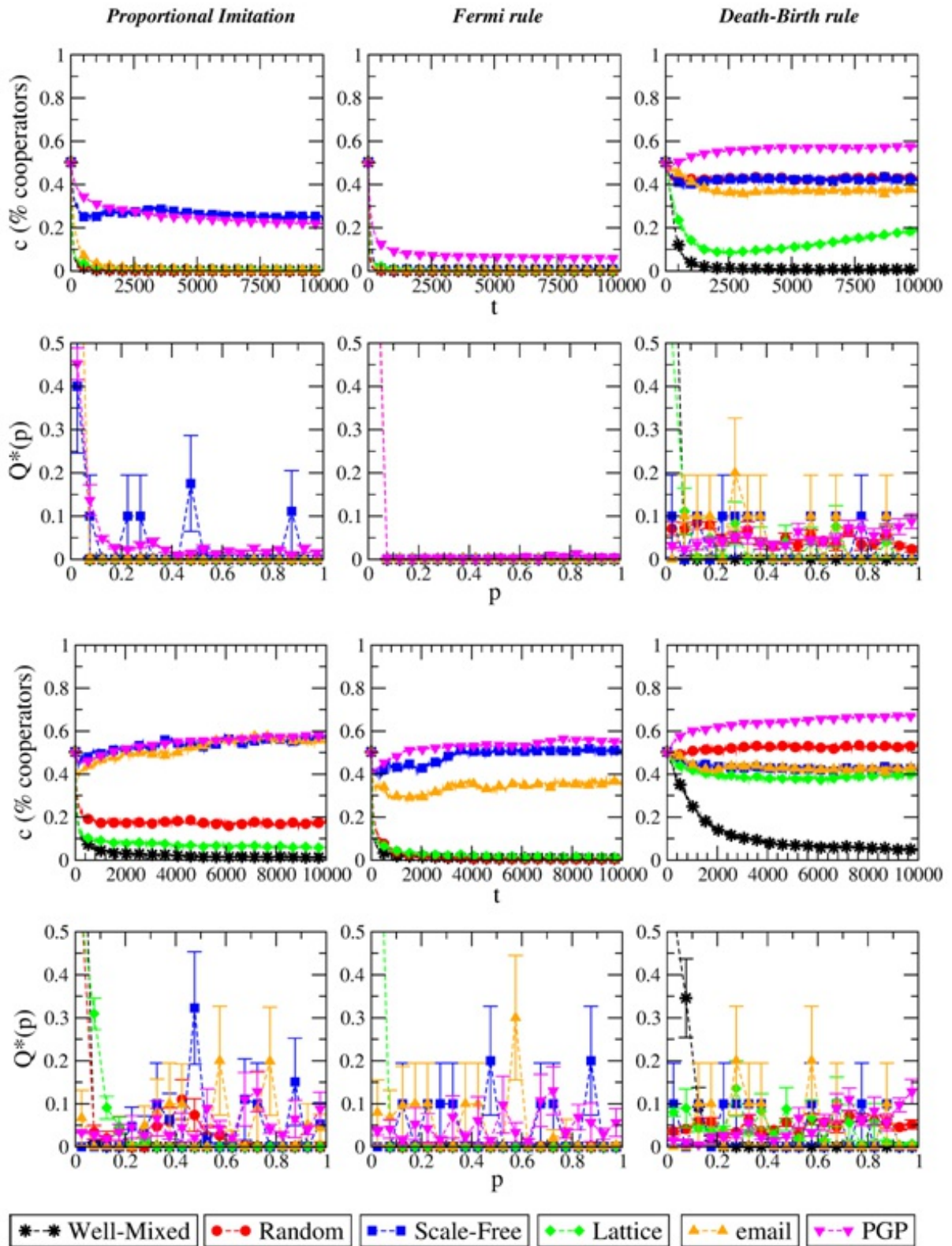


Figure 2. Evolution of the level of cooperation c and stationary distribution of ρ when the evolutionary dynamics is, from left to right: Proportional Imitation, Fermi Rule with $\beta=1/2$, Death-Birth rule. Top plots refer to $S=-1/2$, bottom plots to $S=0$. $T=3/2$ in all cases. Results are averaged over 10 independent realizations.

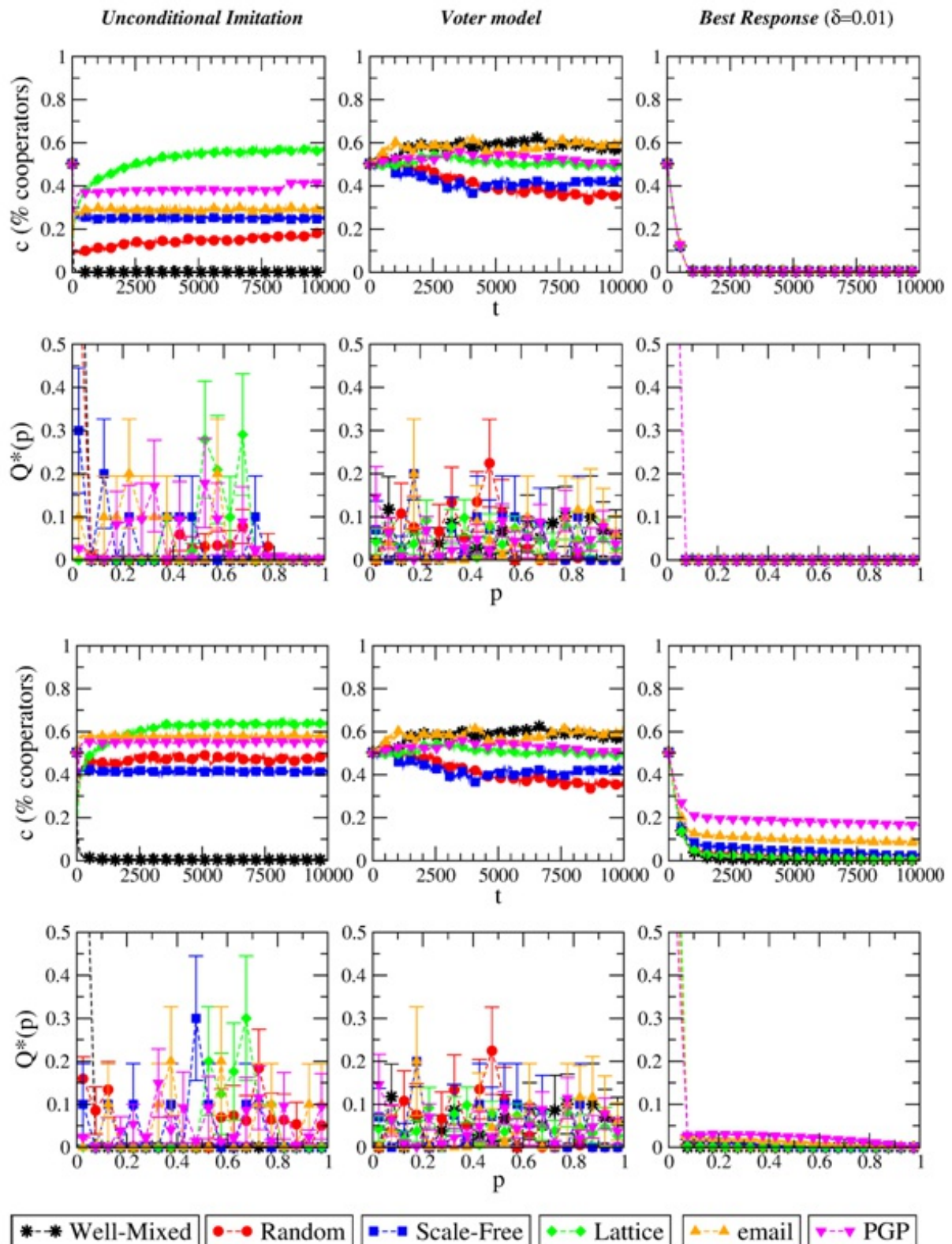


Figure 3. Evolution of the level of cooperation c and stationary distribution of p when the evolutionary dynamics is, from left to right: Unconditional, Voter model, Best Response with $\delta=0.01$. Top plots refer to $S=-1/2$, bottom plots to $S=0$. $T=3/2$ in all cases. Results are averaged over 10 independent realizations.

3.4 While explaining in detail the effects of a particular updating rule and of a given population structure is out of the scope of the present work, we can still gain qualitative insights on the system's behavior from simple observations. Here we discuss the case of networks with highly heterogeneous degree distribution, such as scale-free ones. These topologies are characterized by the presence of players with high degree ("hubs") that generally get higher payoff than an average player's as they play more instances of the game (the average payoff being greater than 0). For a dynamics of imitative nature, hubs' strategy remains stable: they hardly copy their less-earning neighbors, who in turn tend to imitate the hubs. As a result, hubs' strategy spread locally over the network, and, if such strategy profile results in frequent cooperation, a stable subset of player inclined to cooperate can appear around these hubs (Gómez-Gardeñes et al. 2007). The same situation cannot occur in random or regular graphs, where the degree distribution is more homogeneous and there are no hubs with systematic payoff advantage. This phenomenon becomes evident with Proportional Imitation as the updating rule. Note that the subset of players around hubs loses

stability if they can make mistakes (as with the Fermi rule); on the other hand, such stability is enhanced when the updating rule preferentially selects players with high payoffs (as with the Death-Birth rule and Unconditional Imitation), because hubs' strategy spreads more easily. The fact that in the latter two cases cooperation thrives also in lattices is instead related to the emergence of clusters of mutually connected players who often cooperate, get higher payoff than the defectors at the boundary of the cluster exploiting them and can thus survive.

- 3.5 We finally remark that, beyond all the particular features and outcomes of each imitative dynamics, the main conclusion of this analysis is that imitation based on payoff comparison (as modeled by the various processes we consider here) does not lead to the absence of network reciprocity. The only updating rule whose behavior is not affected by the population structure is the Voter model—which however does not depend on payoffs. Such a conclusion is consistent with Yamauchi et al. (2010, 2011), who also considered various imitative dynamics based on payoff comparison and found that the final cooperation level does depend on the network topology (and is in general higher for scale-free networks and lattices). We note also that in Ohtsuki et al. (2006), where the "helping" game^[3] was studied, payoff-based imitation was shown to lead to slightly different situations depending on the network type.

Non-imitative dynamics

- 3.6 The first general remark about these evolutionary rules is that they allow extinct strategies to be reintroduced in the population; because of this, many strategy profiles survive at the end of each realization of the system (even when the dispersion of the parameters p is small). Results for Best Response dynamics are reported in the right column of Figure 3. We recall that this way of updating the strategies is the most "rational" among those we are considering in this work, and is not based on comparing own payoffs with those of others. As a result, we see that for the strong PD the system converges towards full defection for any value of the temptation and for any population structure. Indeed, this outcome is the Nash equilibrium of PD games, which would have been obtained also by global maximization of the individual expected payoffs. Hence the specific value of δ (the amount by which the parameters p can be shifted at each update) only influences the time of convergence to full defection, with higher δ causing simulations to get faster to $p=0 \forall i$. For this reason, we only show results for a particular value of δ . For the weak PD we observe instead a semi-stationary, non-vanishing yet slightly decreasing level of cooperation—which is the consequence of actions C and D bringing to the same payoff when facing a defector. Such cooperation level seems to depend on the network size (bigger networks achieve higher c), rather than on the network topology. In this sense, we can claim that evolution by Best Response features absence of network reciprocity. This conclusion is supported by the fact that the optimal choice for a strong PD game does not depend on what the others do or gain (also for $S=0$ such dependence is only "weak"); as a consequence, the social network in which players are embedded must play no role.
- 3.7 The other non-imitative rule that we consider in this study is Reinforcement Learning. Results for this choice of the dynamics are shown in Figure 4. We start with the simplest assumption of aspiration levels remaining constant over iterations of the game. Here the most remarkable finding is that, in contrast to all other update schemes discussed so far, with this dynamics mixed strategies that are stable attractors of the evolution do appear: the parameters p concentrate around some stationary, non-trivial values, that do not depend on the initial condition of the population, neither on the topology of the underlying network. Concerning cooperation levels, when aspirations are midway between the punishment and reward payoffs ($P < A < R$) we observe a stationary, non vanishing c around $0.3 \div 0.4$ —which is in agreement with what is observed in experiments (Grujic et al. 2010; Gracia-Lázaro et al. 2012b), also with respect to the convergence rate to the stationary state. The specific value of c does depend on the payoff's matrix entries T and S , but not on the population structure. Note that the described behavior is robust with respect to the learning rate λ .^[4] We can thus assert that Reinforcement Learning represents another evolutionary dynamics that can explain the absence of network reciprocity. This happens because, for this choice of updating rule, players do not take into consideration others' actions nor payoffs when adjusting their strategy, and thus the patterns of social interactions become irrelevant. Additional evidence for the robustness of this updating scheme derives from the behavior observed for other aspiration levels, including dynamic ones. As a general remark, the final level of cooperation reached is higher for higher A . For instance, when $R < A < T$ an outcome of mutual cooperation does not meet players' aspirations, however an outcome of mutual defection is far less satisfactory and brings to a substantial increase of p for the next round. Because of this feedback mechanisms, players' strategies tend to concentrate around $p=1/2$ (i.e., playing C or D with equal probability), which thus results in $c=1/2$, again irrespectively of the population structure. Leaving aside the questionable case of aspiration levels below punishment ($S < A < P$), we finally move to the case of adaptive aspiration levels. What we observe now is that players' aspirations become stationary—with final values falling in the range $P < A < R$ —and that no topological effects are present (as with Best Response, bigger networks achieve slightly higher cooperation).

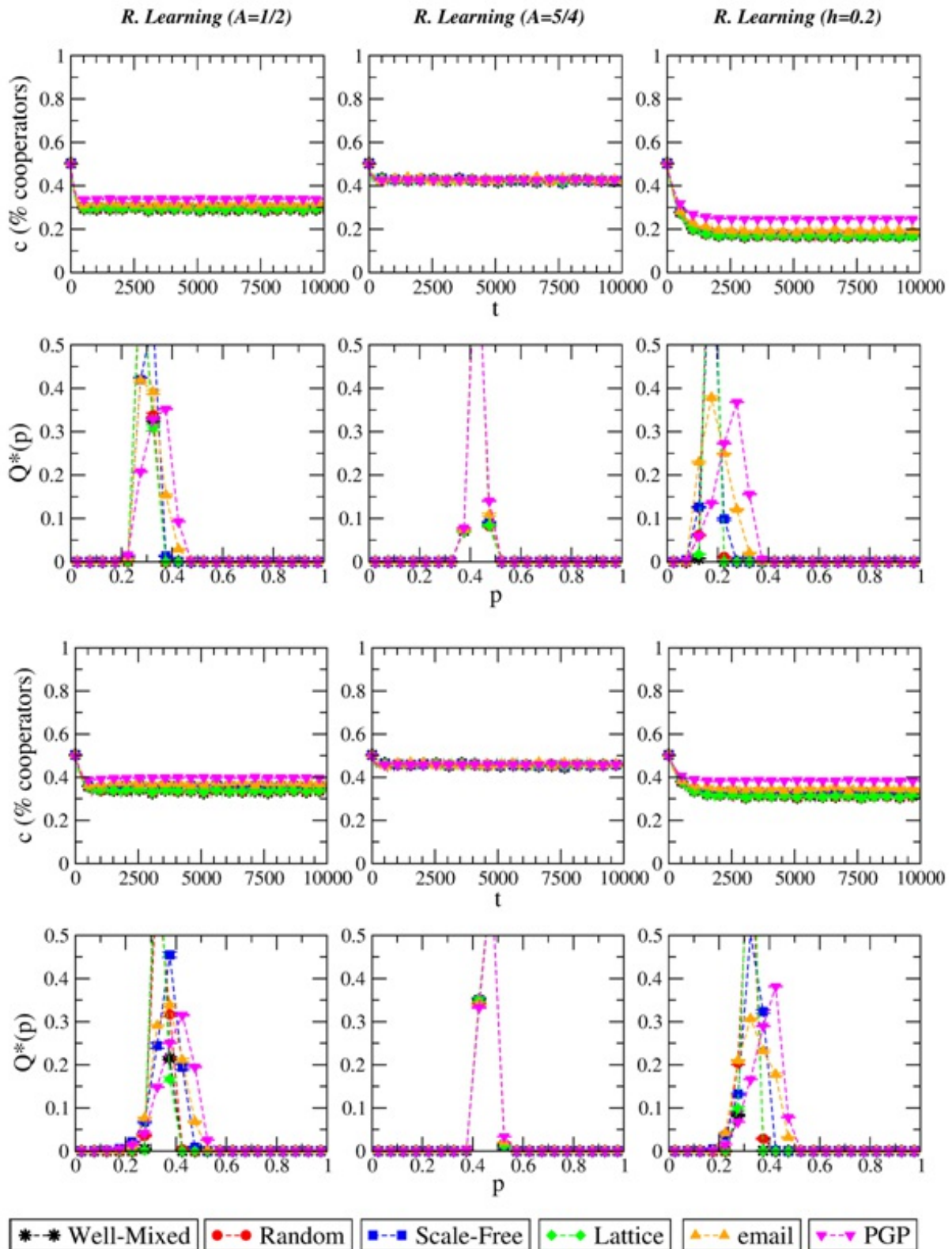


Figure 4. Evolution of the level of cooperation c and stationary distribution of p when the evolutionary dynamics is Reinforcement Learning with $A=0.01$. From left to right: $A=1/2$, $A=5/4$, adaptive A ($A^{(0)}=1/2$, $h=0.2$). Top plots refer to $S=-1/2$, bottom plots to $S=0$. $T=3/2$ in all cases. Results are averaged over 10 independent realizations.

3.8 Summing up, we observe absence of network reciprocity for the innovative dynamics considered here, which we recall are not based on payoff comparison. This again supports our assumption that such outcome derives from not taking into account the payoffs of the rest of the players.

Conclusion

4.1 Understanding cooperation represents one of the biggest challenges of modern science. Indeed, the spreading of cooperation is involved in all major transitions of evolution (Maynard Smith & Szathmary 1995), and the fundamental problems of the modern world (resource depletion, pollution, overpopulation, and climate change) are all characterized by the tensions typical of social

dilemmas. This work has been inspired by the experimental findings (Grujic et al. 2010; Gracia-Lázaro et al. 2012b) that network reciprocity is not a mechanism to promote cooperation within humans playing PD. We aimed at identifying the evolutionary frameworks that are affected by the interaction patterns in the population, thus being inconsistent with the experimental outcomes. To this end, we have considered several mechanisms for players to update their strategy—both of imitative nature and innovative mechanisms, as well as rules based on payoff comparison and others based on non-economic or social factors.^[5] These rules indeed represent most of the alternatives that have been proposed so far in the literature to implement the strategy updating process. We stress that this is a very relevant point, as for the first time to our knowledge we are providing an extensive comparison of payoff-based and non-payoff based evolutionary dynamics on a wide class of networks (representing population structures), and, importantly, within the same model settings. Our research points out that the absence of network reciprocity is a general feature of the evolutionary dynamics (among those considered here) in which players do not base their decisions on others' well being. Note that the evolutionary dynamics that we eliminated as possible responsible for how people behave are difficult to justify for humans and in a social context, because they assume very limited rationality that only allows imitating others. Indeed, analysis of experimental outcomes (Grujic et al. 2014) point out that humans playing PD do not base their decisions on others' payoffs. We thus believe that the present work provides a firm theoretical support for these experimental results, and allows to conclude that many of the evolutionary dynamics, based on payoff comparison, used in theory and in simulations so far simply do not apply to the behavior of human subjects and, therefore, their use should be avoided.

- 4.2 Even so, our findings do not exclude the plausibility of other strategy updating in different contexts. For instance, analytical results with imitative dynamics (Wu et al. 2010) display an agreement with experimental outcomes on dynamical networks (Fehl et al. 2011; Wang et al. 2012). It is also important to stress that our findings here relate to human behavior, whereas, other species could behave differently; thus, in Hol et al. (2013) it has been reported that bacteria improve their cooperation on a spatial structure (and this could arise because of more imitative strategies). Finally we stress that our results are valid for PD games in which players are not aware of the identity of their interacting partners. In this respect, network reciprocity can still be a likely mechanism when people actually know each other (Apicella et al. 2012).



Acknowledgements

This work was supported by the Swiss Natural Science Foundation through grant PBFPR2_145872 and by Ministerio de Economía y Competitividad (Spain) through grant PRODIEVO.



Notes

- ¹ Results for artificial networks are robust with respect to system size and link density.
- ² Unconditional imitation is the only imitative dynamics whose plausibility has not been fully excluded by experiments (Grujic et al. 2014), but only if around 30% of actions are mistaken (i.e., an action different from that of the neighbor with the highest payoff is chosen).
- ³ The helping game shows dynamics similar to PD for some choice of parameters (Nowak and Sigmund 1998).
- ⁴ The specific value of λ influences the convergence time of the system to its stationary state. However, if players learn too quickly ($\lambda \approx 1$) then the parameters p change too much at each iteration to reach stationary values, as typically happens for learning algorithms.
- ⁵ Note that we have not considered noise (i.e., the possibility of making errors) in the strategy updating process. Indeed, noise can make the model more realistic (and even eliminate some stable outcomes); however, it can hinder our understanding of the joint effect of network topology and evolutionary dynamics on the system.



References

- APICELLA, C. L., Marlowe, F. W., Fowler, J. H. & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, *481*, 497–501.
- AXELROD, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- BLUME, L. E. (1993). The statistical mechanics of strategic interaction. *Games and Economic Behavior*, *5*, 387–424.
- BOGUÑÁ, M., Pastor-Satorras, R., Diaz-Guilera, A. & Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, *70*: 056122.

- BUSH, R. & Mosteller, F. (1955). *Stochastic Models of Learning*. New York: John Wiley & Son.
- CASSAR, A. (2007). Coordination and cooperation in local, random and small world networks: Experimental evidence. *Games and Economic Behavior*, *58*, 209–230.
- CIMINI, G. & Sánchez, A. (2014). Learning dynamics explains human behavior in Prisoner's Dilemma on networks. *Journal of the Royal Society: Interface*, *11*: 94 20131186.
- DAWES, R. M. (1980). Social Dilemmas. *Annual Review of Psychology*, *31*, 169–193.
- FEHL, K., van der Post, D. J. & Semmann, D. (2011). Co-evolution of behaviour and social network structure promotes human cooperation. *Ecology Letters*, *14*, 546–551.
- FEHR, E. & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*, 159–181.
- FISCHBACHER, U., Gächter, S. & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a Public Goods experiment. *Economic Letters*, *71*: 397404.
- GÓMEZ-GARDEÑES, J., Campillo, M., Floria, L. M. & Moreno, Y. (2007). Dynamical organization of cooperation in complex topologies. *Physical Review Letters*, *98*: 108103.
- GRACIA-LÁZARO, C., Cuesta, J., Sánchez, A. & Moreno, Y. (2012a). Human behavior in Prisoner's Dilemma experiments suppresses network reciprocity. *Scientific Reports*, *2*, 325.
- GRACIA-LÁZARO, C., Ferrer, A., Ruiz, G., Tarancón, A., Cuesta, J., Sánchez, A. & Moreno, Y. (2012b). Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. *Proceedings of the National Academy of Sciences U.S.A.*, *109*, 10922–20926.
- GRUJIC, J., Fosco, C., Araujo, L., Cuesta, J. & Sánchez, A. (2010). Social experiments in the mesoscale: Humans playing a spatial Prisoner's Dilemma. *PLoS ONE*, *5*: e13749.
- GRUJIC, J., Gracia-Lázaro, C., Milinski, M., Semmann, D., Traulsen, A., Cuesta, J., Moreno, Y. & Sánchez, A. (2014). A comparative analysis of spatial Prisoner's Dilemma experiments: Conditional cooperation and payoff irrelevance. *Scientific Reports*, *4*, 4615.
- GUIMERÀ, R., Danon, L., Diaz-Guilera, A., Giralt F. & Arenas A. (2003). Self-similar community structure in a network of human interactions. *Physical Review E*, *68*: 065103(R).
- HARDIN, G. (1968). The Tragedy of the Commons. *Science*, *162*, 1243–1248.
- HAUERT, C. & Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the Snowdrift game. *Nature*, *428*, 643–646.
- HELBING, D. (1992). Interrelations between stochastic equations for systems with pair interactions. *Physica A*, *181*, 29–52.
- HELBING, D. & Yu, W. (2010). The future of social experimenting. *Proceedings of the National Academy of Sciences U.S.A.*, *107*, 5265–5266.
- HOL, F. J. H., Galajda, P., Nagy, K., Woolthuis, R. G., Dekker, C. & Keymer, J. E. (2013). Spatial structure facilitates cooperation in a Social Dilemma: Empirical evidence from a bacterial community. *PLoS ONE*, *8*: e77402.
- HOLLEY, R. & Liggett, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the Voter Model. *The Annals of Probability*, *3*, 643–663.
- HOFBAUER, J. & Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press.
- HOFBAUER, J. & Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, *40*, 479–519.
- IZQUIERDO, S. S., Izquierdo, L. R. & Gotts, N. M. (2008). Reinforcement Learning dynamics in social dilemmas. *Journal of Artificial Societies and Social Simulation*, *11* (2): 1: <http://jasss.soc.surrey.ac.uk/11/2/1.html>.
- KIRCHKAMP, O. & Nagel, R. (2007). Naive learning and cooperation in network experiments. *Games and Economic Behavior*, *58*, 269–292.
- MACY, M. W. & Flache, A. (2002). Learning dynamics in social dilemmas. *Proceedings of the National Academy of Sciences U.S.A.*, *99*, 7229–7236.
- MATSUI, A. (1992). Best Response dynamics and socially stable strategies. *Journal of Economic Theory*, *57*, 343–362.
- MAYNARD SMITH, J. & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*, 15–18.
- MAYNARD SMITH, J. & Szathmary, E. (1995). *The Major Transitions in Evolution*. Oxford: Freeman.

- MORAN, P. A. P. (1962). *The Statistical Processes of Evolutionary Theory*. Oxford: Clarendon Press.
- NOWAK, M. A. & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, *359*, 826–829.
- NOWAK, M. A. & Sigmund, K. (1998). The dynamics of indirect reciprocity. *Journal of Theoretical Biology*, *194*, 561–574.
- NOWAK, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*, 1560–1563.
- OHTSUKI, H., Hauert, C., Lieberman, E. & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, *441*, 502–505.
- ROCA, C. P., Cuesta, J. & Sánchez, A. (2009a). Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of Life Reviews*, *6*, 208–249.
- ROCA, C. P., Cuesta, J. & Sánchez, A. (2009b). On the effect of spatial structure on the emergence of cooperation. *Physical Review E*, *80*, 46106.
- ROCA, C. P., Cuesta, J. & Sánchez, A. (2009c). Promotion of cooperation on networks? The myopic Best Response case. *European Physical Journal B*, *71*, 587–595.
- SCHLAG, K. H. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, *78*, 130–156.
- SIGMUND, K. (2010). *The Calculus of Selfishness*. Princeton: Princeton University Press.
- SURI, S. & Watts, D. J. (2011). Cooperation and contagion in web-based, networked Public Goods experiments. *PLoS ONE*, *6*(3): e16836.
- SYSI-AHO, M., Saramäki, J., Kertész, J. & Kaski, K. (2005). Spatial Snowdrift game with myopic agents. *European Physical Journal B*, *44*, 129–135.
- SZABÓ, G. & Toke, C. (1998). Evolutionary Prisoner's Dilemma game on a square lattice. *Physical Review E*, *58*, 69–73.
- SZABÓ, G. & Fáth, G. (2007). Evolutionary games on graphs. *Physics Reports*, *446*, 97–216.
- TRAULSEN, A., Semmann, D., Sommerfeld, R. D., Krambeck, H. J. & Milinski, M. (2010). Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences U.S.A.*, *107*, 2962–2966.
- VILONE, D., Ramasco, J. J., Sánchez, A. & San Miguel, M. (2012). Social and strategic imitation: The way to consensus. *Scientific Reports*, *2*, 686.
- VILONE, D., Ramasco, J. J., Sánchez, A. & San Miguel, M. (2014). Social imitation vs strategic choice, or consensus vs cooperation in the networked Prisoner's Dilemma. *Physical Review E*, *90*, 022810.
- WANG, J., Suri, S. & Watts, D. J. (2012). Cooperation and assortativity with dynamic partner updating. *Proceedings of the National Academy of Sciences U.S.A.*, *109*, 14363–14368.
- WU, B., Zhou, D., Fu, F., Luo, Q., Wang, L. & Traulsen, A. (2010). Evolution of Cooperation on Stochastic Dynamical Networks. *PLoS ONE*, *5*(6): e11187.
- YAMAUCHI, A., Tanimoto, J. & Hagushima, A. (2010). What controls network reciprocity in the Prisoner's Dilemma game? *Biosystems*, *102*(2–3), 82–87.
- YAMAUCHI, A., Tanimoto, J. & Hagushima, A. (2011). An analysis of network reciprocity in Prisoner's Dilemma games using full factorial designs of experiment. *Biosystems*, *103*(1), 85–92.