

# BLIND VS. EMBEDDED INDIRECT RECIPROCITY AND THE EVOLUTION OF COOPERATION

Simone Righi

Department of Agricultural and Food Sciences – University of Bologna

I-40038, Bologna, Italia

Email: s.righi@unibo.it

and

“Lendület” Research Center for Education and Network Studies

Hungarian Academy of Sciences, Centre for Social Sciences

H-1014, Budapest, Hungary

Karoly Takacs

“Lendület” Research Center for Education and Network Studies

Hungarian Academy of Sciences, Centre for Social Sciences

H-1014, Budapest, Hungary

Email: takacs.karoly@tk.mta.hu

## KEYWORDS

Agent-based modelling; Evolution of Cooperation;  
Indirect reciprocity; Forgiveness; Speed of evolution

## ABSTRACT

The evolution of cooperation is one of the fundamental problems of both social sciences and biology. It is difficult to explain how a large extent of cooperation could evolve if individual free riding always provides higher benefits and chances of survival. In absence of direct reciprocation, it has been suggested that indirect reciprocity could potentially solve the problem of large scale cooperation. In this paper, we compare the chances of two forms of indirect reciprocity with each other: a blind one that rewards any partner who did good to previous partners, and an embedded one that conditions cooperation on good acts towards common acquaintances. We show that these two versions of indirect reciprocal strategies are not very different from each other in their efficiency. We also demonstrate that their success very much relies on the speed of evolution: their chances for survival are only present if evolutionary updates are not frequent. Robustness tests are provided for various forms of biases.

## INTRODUCTION

The evolution of cooperative behaviour within a community of individuals is a largely studied problem (Hoffmann 2000; Sachs et al. 2004). The presence of kinship relationships can explain cooperative behaviour as the willingness to see related individuals - sharing a part of the genetic code - thrive and reproduce (Hamilton 1964; Grafen 1984). More difficult to justify is the cooperation among unrelated individuals.

Direct reciprocity is known to be a powerful mechanism for the evolution of cooperation (Axelrod and Hamilton (1981); Axelrod (1984)). Direct reciprocity, however, is difficult to justify as the only mechanism behind high levels of cooperation in human societies, as its efficiency strongly relies on direct knowledge and perfect memory of past behaviour of interacting partners. Previous research has uncovered various mechanisms that can contribute to the establishment of cooperation. Among others, these mechanisms include constrained interaction on networks or spatial structures (Hauert and Doebeli 2004; Lieberman et al. 2005), the competition among communities (Gunthorsdottir and Rapoport 2006; West et al. 2007; Puurtinen and Mappes 2009), the interaction within small populations (Dunbar 1992), the presence of negative relationships among agents on the side of positive ones (Righi and Takács 2014) and more in general the presence of information about peers coming through mechanisms such as social comparison (Whitaker et al. 2016) reputation (Sigmund 2012), image-scoring (Wedekind and Milinski 2000), gossip (Sommerfeld et al. 2007; Wu et al. 2015), and language (Smith 2010).

Most of these mechanisms are related to the idea of indirect reciprocity: help - or retaliation - does not come from the interacting partners but rather from some third individual (Boyd and Richerson 1989; Nowak and Sigmund 1998, 2005). When interactions are constrained on a network structure (Hauert and Doebeli 2004), there can be at least two different operationalizations of the concept of indirect reciprocity, depending on the assumptions made about the informational flow accessible to agents.

A first type of indirect reciprocity is one where an

individual can observe the behaviour of the interacting partners with any third individual in the population related to the latter. In this case it is assumed that information can freely flow on the network of interactions so that a joint connection between the indirectly reciprocal agents is not necessary for the information about the behaviour of common peers to pass from the former to the latter. Within this paper we call this blind type of operationalization of the indirect reciprocity concept “unconnected reciprocity” (or UR).

A second type of indirect reciprocity is only concerned with the flow of information coming from peers. Within this paper we call this embedded type of operationalization of the indirect reciprocity concept “connected reciprocity” (or CR).

As this two types of indirect reciprocity rely in different information sets, their strength in sustaining cooperation can differ according to the setup studied. The objective and the innovation of this study is to characterize and compare the effectiveness of these two types of indirect reciprocity strategies. We analyze the chances of the two types of indirect reciprocity to support the evolution of cooperation under different conditions, in particular under network dynamics, in relation to the extent of forgiveness, and to the speed of strategy evolution.

## THE MODEL

We consider a model with  $N$  agents placed on a non-weighted and non-directed Erdős-Rényi graph with a given density  $d$ . Every agent  $i \in N$  is characterized by a strategy type and by a set of connections with a subset of the whole population  $F_i^t \subset N$ . Each time step, the two-person single-shot Prisoner’s Dilemma (PD) is played once by each individual with each of her network connections. The agents play the PD, characterized by the classical payoff structure  $Temptation(T) > Reward(R) > Punishment(P) > Sucker(S)$ , with each of their neighbors. Specifically, for the sake of the simulations hereby discussed we fix the value of the payoffs to  $T = 5$ ,  $R = 3$ ,  $P = 1$  and  $S = 0$  as in Axelrod (1984). We consider the following four strategies:

- Unconditional Defection (UD): this type of agent always defects regardless of the behaviour or characteristics of the interaction partners.
- Unconditional Cooperation (UC): this type of agent always cooperates regardless of the behaviour or characteristics of the interaction partners.
- Unconnected Reciprocity (UR): this type of agent reciprocates the last action of the interacting partner with a randomly selected connection of the latter (i.e. some  $z \in F_j^t$ ).<sup>1</sup>
- Connected Reciprocity (CR): this type of agent reciprocates the last action of the interacting partner with a randomly selected common connection (i.e. some  $z \in F_i^t \wedge F_j^t$ ).<sup>2</sup>

<sup>1</sup>If there is no previous action to observe than the UR acts randomly, with 50% chance of cooperation.

<sup>2</sup>If there is no previous action to observe than the CR acts randomly, with 50% chance of cooperation.

During every dyadic interaction, agents can observe the past behaviour of the partner with all his connections and play according to their type. One key difference between the CR and the UR strategy is that the former can only get information to condition its behaviour from closed triads, while the latter can also condition its choice on the behaviour of individuals unconnected with him.

Time is divided in discrete periods and simulations run until a stable equilibrium is reached or 10000 periods have passed. At each time step  $t$ , each agent  $i$  contemporaneously plays the PD with all agents his first order social neighborhood, i.e. with each  $j \in F_i$  observing all actions performed by others in  $t - 1$ . Agents of type UR and CR, when observing their partners’ actions may observe defection; in this case –with probability  $P_{for}$ – they may decide to give the partner another opportunity and observe another of his actions. This simulates the fact that individuals may forgive a defection act. At the end of the interaction phase, each individual computes his average payoff from all his interactions and compares it with that of peers in the direct neighborhood (i.e. with all those he has played with). With probability  $P_{evo}$  the individual changes its strategy into the one of the best performing partner.<sup>3</sup> Hence, the evolutionary strategy we apply is the “copy the best” update rule.

**Network dynamics.** Following hints of previous research (Santos et al. 2006), network dynamics may provide favorable conditions for the evolution of cooperation. For this reason it is important to observe its impact on the effectiveness of indirect reciprocity strategies on cooperation. In this model, each of ties of each individuals  $F_i^t$  can be subject to rewiring, with some probability at each time step. We consider three mechanisms of network update.

- With probability  $P_{ntw}$  a selected connection is rewired to another agent selected uniformly at random from the population.
- With probability  $P_{open}$  a selected connection is rewired so to open a closed triad incident on that particular edge.
- With probability  $P_{close}$  the rewiring is done by closing a triad that was open.

The three probabilities are extracted independently – for each link – at each time step. For this reason the sign of the difference  $P_{open} - P_{close}$  indicates the propensity of the network clustering to increase (negative signs of the difference) or decrease (positive signs of the difference) over time.

**Outcome Measures.** We explore our model running a large number of simulations (1000 for the simulations proposed here) and by studying the statistical relationship between the outcome variables and the various parameters. This type of analysis is chosen to explore a complex model with many different parameters extensively. As outcome variables, the number of CR and the number of UR strategies are used alongside

<sup>3</sup>In case of ties, one of the partners with the highest payoff is selected randomly

with the proportion of cooperative acts. Populations of 200 agents are studied for a large number of time periods (10,000) or until the whole population of agents becomes of the same type so that no further evolution is possible. We define this latter moment  $t_{conv}$ . Since the model characteristics do not guarantee the convergence of the population to a single strategy, to describe evolutionary success, we define a strategy to gain

- **Absolute dominance:** if each individual in the population ends up playing the strategy or if – when  $t_{max} = 10000$  is reached – this strategy has been adopted by more than 90% of the individuals.
- **Relative dominance:** if  $t_{max} = 10,000$  is reached and this strategy is adopted by the relative majority of individuals.

Obviously, the second mode of dominance should be computed only for those cases where no convergence to a single strategy has been achieved. Most of our results show a quasi-certain convergence on one type of strategy. Therefore, except when otherwise noted, we will refer to the concept of absolute dominance in the discussion of our results.

In addition, the proportion of cooperative acts is computed at  $t_{conv}$  or  $t_{max}$  (in case of non-convergence).

## RESULTS

**Indirect reciprocity strategies and cooperation.** We start by running 1000 simulations considering a population that is initially equally divided between the four types of strategies: UC, UD, CR and UR.  $N = 200$  agents are laid on a random network with density  $d = 0.2$ .<sup>4</sup> For each single simulation the values of  $P_{evo}$  (the strategy evolution probability),  $P_{ntw}$  (probability of random network change),  $P_{for}$  (probability of forgiveness),  $P_{open}$  (probability to open each triad at each step) and  $P_{close}$  (probability to close each triad at each step) are extracted uniformly and independently from uniform distributions between 0 and 1.

Figure 1 shows that cooperation rarely disappears completely. On the opposite, for the majority of the simulations, the final proportion of cooperators is above zero. This implies that, for the largest part of possible parameter combinations, unconditional defection is unable to completely eliminate other types of strategies. The proportion of simulations with a share of cooperating acts larger than 1/2, however, is smaller, indicating that cooperation seldom becomes the action adopted by the majority of agents. It is interesting to note the sharp drop in the frequency of cooperation around a proportion 1/2, indicating that while the strategies of indirect reciprocity survive, cooperation seldom gains relative dominance.

Table I reports that UD was dominant in 83% of the simulations in which a strategy gained absolute dominance. As expected, UC never gained absolute dominance. More remarkable is the significant presence of

<sup>4</sup>The size and density of the network are calibrated throughout the article to resemble that of human ancestors' communities, in line with Dunbar (1992). Moreover from the technical viewpoint the relatively high density almost surely ensures the absence of isolates.

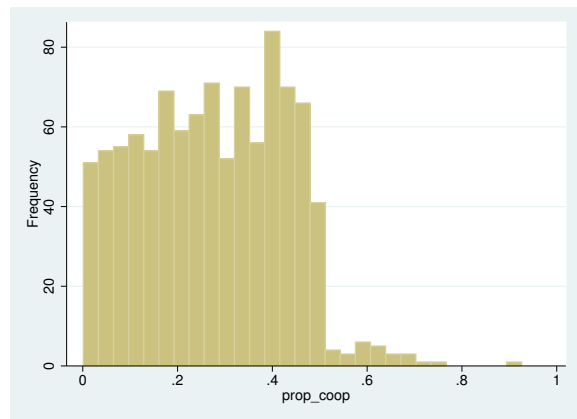


Fig. 1. Distribution of proportion of cooperation actions at  $t_{conv}$ . Results from 1000 simulations with populations initially equally divided among the four types of strategies. The value of  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$  and  $P_{close}$  are selected randomly and independently from a distribution between 0 and 1.

instances in which CR and UR became dominant in the simulations that resulted in a single winner strategy. Indeed, about 10% of these runs ended up with the dominance of the CR strategy, while 7% of them terminated with the win of the UR strategy. Since the parameter space explored is quite large and contains parameter combinations very hostile to the emergence of cooperation, the relatively small success of indirect reciprocity strategies is noteworthy. Furthermore, the high standard deviation observed around the average number of individuals adopting a strategy at the end of a simulation indicates the potential presence of very different outcomes.

**Blind indirect reciprocity and cooperation.** In order to understand the comparative strength of blind and embedded indirect reciprocal strategies, it is convenient to study separately the cases where only one of them is present alongside the baseline strategies of UD and UC. We run 1000 simulations with initial populations equally divided among UD, UC and UR and with  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  extracted uniformly and independently from uniform distributions between 0 and 1. As Figure 2 (Left Panel) shows, the final number of URs has a bimodal distribution. In many cases, UR disappears and the population is progressively dominated by UD. In others, the UR strategy becomes the absolutely dominant one, spreading to the whole population. Simulations characterized by the survival of UR as a non dominant strategy are very rare. The Left Panel of Figure 2 (Left Panel) explores the outcomes for different probabilities of forgiveness  $P_{for}$  and clearly shows that UD has the potential to become the absolute dominant strategy under any level of forgiveness. The Right Panel of Figure 2 (Right Panel) shows how the proportion of cooperation increases in cases dominated by the UR strategy with higher levels of forgiveness  $P_{for}$ . Notice that not all cases where UR gains dominance are characterized by cooperative behavior. Blind indirect reciprocity often leaves a blind eye for an eye.

Strategy type	%Abs. Dom.	Avg. num.	Std Num	min	max
UD	83%	115.59	98.68	0	200
UC	0 %	0.12	1.35	0	26
CR	10%	13.49	49.40	0	200
UR	7%	10.26	43.31	0	200

TABLE I: Characteristics of simulations absolutely dominated by a strategy. *Notes:* Columns show the proportion of these runs in which each strategy gained absolute dominance, the average number, the standard deviation, the minimum, and the maximum number of agents playing that strategy across these simulation runs. All simulations start from a population equally divided among the four types of strategies. The values of  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$  and  $P_{close}$  are selected randomly and independently from a distribution between 0 and 1.

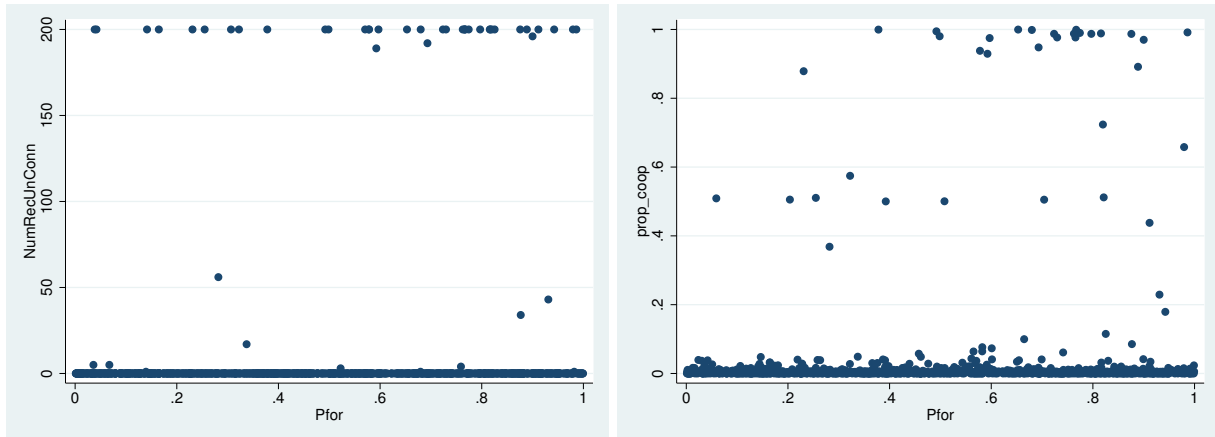


Fig. 2. Left Panel: Number of URs at the end of the simulation. Right Panel: Proportion of cooperative actions at the end of the simulation. Both panels represent the relevant variable against the level of  $P_{for}$  associated with the simulation. In all 1000 shown simulations the population is initially equally divided between UD, UC and UR strategies, while  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  are extracted uniformly at random between 0 and 1.

Let us now study the relationship between the perfectness or the speed of evolution and the chances of blind indirect reciprocity. We find that faster evolutionary dynamics is associated with lower chances for the survival of cooperative behaviour. Indeed, the Left Panel of Figure 3 shows that fast evolution completely eliminates the UR strategy: when the speed of evolution is too high the UR strategy cannot adapt fast enough to the defecting behaviour of UDs and disappears consequentially. Observing the Right Panel of Figure 3 it is immediately evident that the maximum level of cooperation achievable by UR strategies decreases in the speed of evolution, after the latter starts to exceed the threshold  $P_{evo} = 0.2$ . Above this threshold the strategy is unable to sustain the absolute dominance of cooperation.<sup>5</sup>

To conclude our analysis of UR strategies, we now study statistically the level of dominance of this strategy in relation to our key parameter values. For this, we discretize the number of URs at the end of the simulation creating a variable  $catUR$  that is assigned value 1 if the population of UR strategies is larger than 100, and value of 0 otherwise. We then run a probit regression using  $catUR$  as the dependent variable and the value of the parameters  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  as independent variables. Results are reported

<sup>5</sup>The value of 1/2 at  $P_{evo} = 1$  is due to the constant updating of strategies that start with a 50% probability of cooperation.

catUR	Coef.	Std. Err.	$z$	$P >  z $
$P_{evo}$	-1.351303	.3076922	-4.39	0.000
$P_{ntw}$	-.4261091	.2804871	-1.52	0.129
$P_{for}$	.6986143	.2758046	-2.53	0.011
$P_{open}$	-.2280673	.2728396	0.84	0.403
$P_{close}$	-.1347793	.2696162	-0.50	0.617
Const.	-1.451655	.3154203	-4.60	0.000

TABLE II: Probit regression of the number of URs at the end of each simulation (discretized as discussed in the main text) on key parameter values.

in Table II.

Table II shows that, as suggested qualitatively by Figure 3,  $P_{evo}$  has a strong negative effect (with a p-value lower than 0.01) on the chances of UR to become dominant. On the contrary – and again in line with our qualitative findings, – forgiveness has a positive and significant effect on the chances of UR to become dominant. Finally,  $P_{ntw}$ ,  $P_{open}$  and  $P_{close}$  have non-significant coefficients. This implies that the variations in the parameters of network dynamics do not affect significantly the outcomes discussed. The lack of significance of these parameters can be interpreted as a sign of the robustness of the conclusions reached on the conditions that sustain the emergence of cooperation through blind indirect reciprocity.

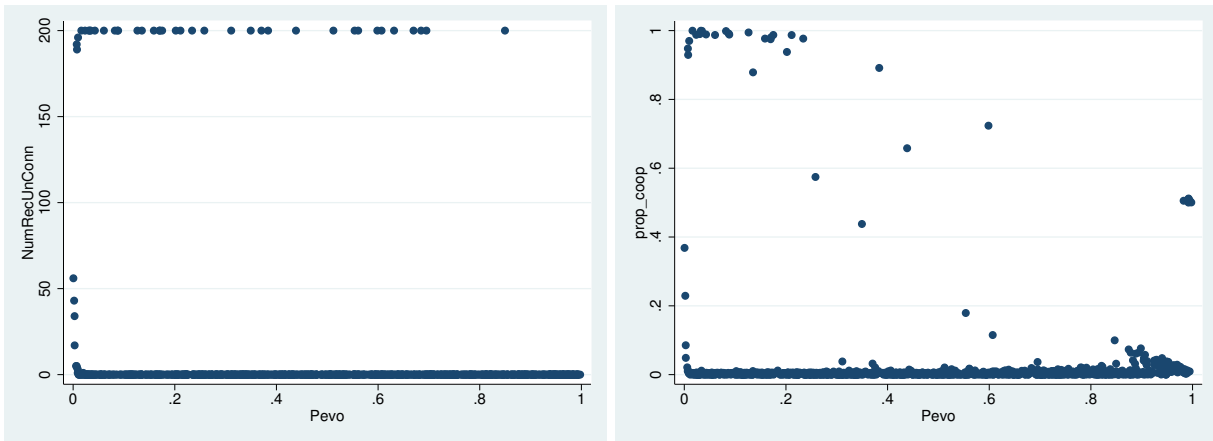


Fig. 3. Left Panel: Number of URs at the end of the simulation. Right Panel: Proportion of cooperative actions at the end of the simulation. Both panels represent the relevant variable against the level of  $P_{evo}$  associated with the simulation. In all 1000 shown simulations the population is initially equally divided between UD, UC and UR strategies while  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  are extracted uniformly divided between 0 and 1.

**Embedded indirect reciprocity and cooperation.** We run a set of 1000 simulations with initial populations equally divided among UDs, UCs and CRs.  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  are all extracted uniformly and independently from uniform distributions between 0 and 1. As Figure 4 (Left Panel) shows, the general relationship between the number of CR strategy registered at the end of the simulation and  $P_{for}$  is in line with the one observed in Figure 2 for UR: the majority of simulations ended up with the dominance of UD, but some with the dominance of the CR strategy. The Right Panel of Figure 4 highlights, however, an interesting difference concerning the levels of cooperation that can be sustained. For low levels of the  $P_{for}$  parameter, while CR can become dominant, it does so by mimicking the UD strategy, thus the whole population acts as defectors. Only when  $P_{for} > 0.4$ , cooperation grows significantly, eventually reaching absolute dominance. This is in contrast with what is observed for UR that sustains high levels of cooperation also for lower levels of forgiveness. This difference is due to the underlying difference between the CR and the UR strategies. Indeed, blind indirect reciprocity (UR) has potentially access to a larger number of individuals as a potential source of information for conditioning its behaviour toward the interacting partner. On the contrary, CR is restricted to interact with common partners ( $F_j^t \wedge F_i^t$ ), and hence forgiveness provides less additional advantages for its success.

Figure 5 shows that the speed of evolution is not favorable for the embedded indirect reciprocity strategy. Despite the broad similarity with Figure 3, there are two notable differences. First, CR can become dominant somewhat more likely than UR with fast strategy evolution. Second, the number of simulations in which a high level of cooperation is sustained is somewhat higher in Figure 5, but drops to zero sharply around  $P_{evo} = 0.4$ . In contrast, Figure 3 displays a progressive decrease by the speed of evolution, but with more internal variation. The reason for this slight difference is - again - the different set of information accessible

catCR	Coef.	Std. Err.	$z$	$P >  z $
$P_{evo}$	-1.144433	.3030948	-3.78	0.000
$P_{ntw}$	.1400632	.2657115	-0.53	0.598
$P_{for}$	.9659768	.2868299	-3.37	0.001
$P_{open}$	.3467598	.2707578	1.28	0.200
$P_{close}$	-.0308959	.2661878	-0.12	0.908
Const.	-2.073514	.3350006	-6.19	0.000

TABLE III: Probit regression of the number of CRs at the end of each simulation (discretized as discussed in the main text) on parameter values assumed by each of the model parameters.

to the two indirectly reciprocal strategies. The results show that the larger, but sparser, network of information accessible to UR allows for faster adaptation to defecting behaviour than the tight but smaller network of CR, thus allowing for higher levels of cooperation to be sustained for  $0.4 < P_{evo} < 1$ .

We finally run a statistical analysis of the level of dominance of CR in relation to key parameter values. Similarly to the UR case, we discretized the number of CRs at the end of the simulation creating a variable  $catCR$  that - for each simulation - takes value 1 if the number of CR is larger than 100 and value 0 otherwise. We then run a *probit* regression using  $catCR$  as dependent variable and parameters  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  as independent variables.

Results are reported in Table III which confirms our qualitative results also for CR. The chances of CR to become dominant are shown to decrease significantly in  $P_{evo}$  and significantly increase in  $P_{for}$ . Both results have very high levels of significance with  $p \leq 0.001$ .

Finally, the parameters associated with  $P_{close}$ ,  $P_{open}$ , and  $P_{ntw}$  are not significant. This is reassuring about the robustness of our findings under different parametrizations of the network dynamics.

## CONCLUSIONS

Against rational interests and equilibrium predictions, humans display a high extent of cooperation, also

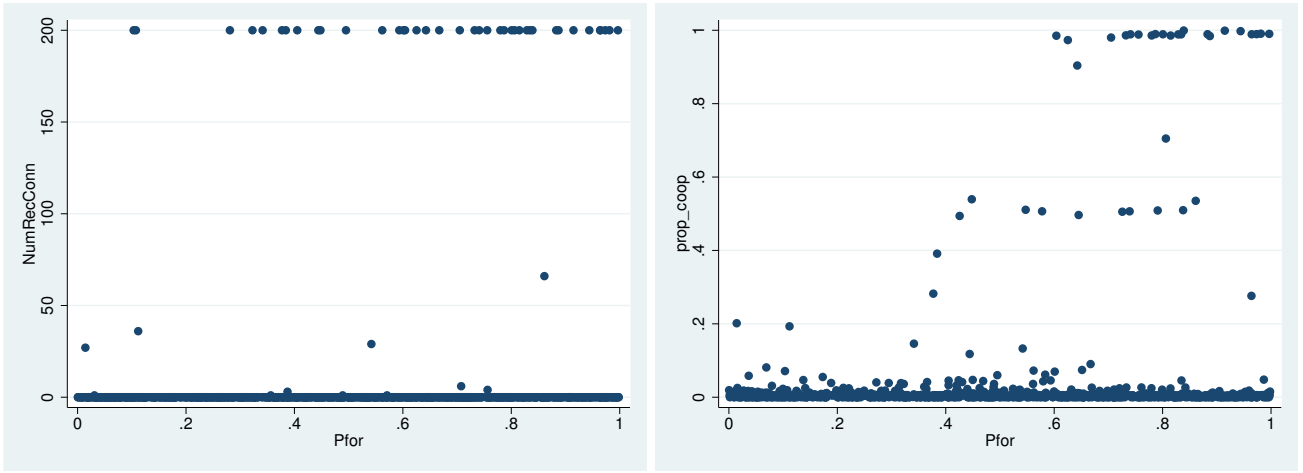


Fig. 4. Left Panel: Number of CRs at the end of the simulation. Right Panel: Proportion of cooperative actions at the end of the simulation. Both panels represent the relevant variable against the level of  $P_{for}$  associated with the simulation. In all 1000 simulations the population is initially equally divided between UD, UC and CR strategies while  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  are extracted uniformly at random between 0 and 1.

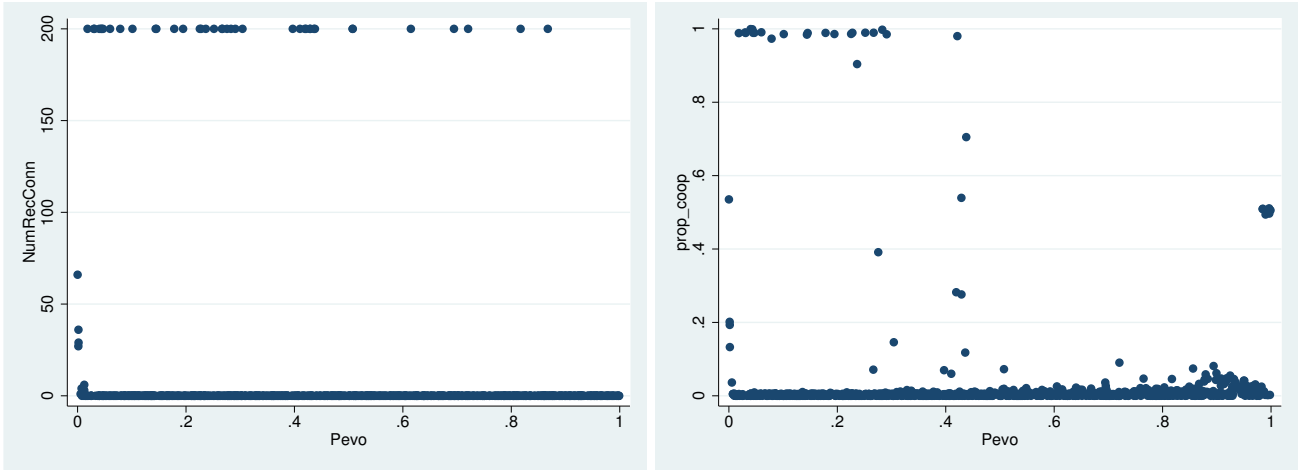


Fig. 5. Left Panel: Number of CRs at the end of the simulation. Right Panel: Proportion of cooperative actions at  $t_{conv}$  or at  $t_{max}$ . Both panels represent the relevant variable against the level of  $P_{evo}$  associated with the simulation. In all 1000 simulations the population is initially equally divided between UD, UC and CR strategies while  $P_{evo}$ ,  $P_{ntw}$ ,  $P_{for}$ ,  $P_{open}$ , and  $P_{close}$  extracted uniformly at random between 0 and 1.

against strangers. Previous research has suggested that indirect reciprocity might explain substantial cooperation levels. The objective of this study was to assess the effectiveness of two different types of indirect reciprocity strategies in their capacity to sustain cooperation. We modeled interactions between agents with the two-person Prisoner’s Dilemma (PD), which is the most puzzling and best studied social dilemma situation. In our model, agents represented humans interacting on a random and developing network. We implemented three mechanisms of network change and examined the chances of cooperation by two different types of indirect reciprocity strategies under a full range of combinations. Precisely, we assumed different combinations of random rewiring of ties, a mechanism of opening triads (creating structural holes), and a mechanism of closing triads (increasing clustering).

Network relations were crucial for our model, because these connections provided agents with the opportunity to condition their actions on the information they re-

ceived from and about network partners. We studied two forms of indirect reciprocity: a blind one that rewards any partner who did good to previous partners, and an embedded one that conditions cooperation on good acts towards common acquaintances.

Our results show that both strategies sustain similar levels of cooperation in situations characterized by the presence of unconditional defectors. Moreover, for both strategies their success depend strongly and negatively on the speed of evolution for strategies. In contexts characterized by fast evolution, both strategies are dominated by defection and disappear or become functionally equivalent to unconditional defectors, thus surviving by always defecting. Forgiveness – the probability of indirectly reciprocal strategies to give a second opportunity to defecting partners by observing their behaviour with a second partner – is shown to have a positive effect on the capacity of both strategies to sustain cooperation.

While the two types of indirect reciprocity strategies

share similar patterns of behaviour, our analysis also uncovered differences in their efficiency. First, the blind indirectly reciprocal strategy (UR) supports relatively higher levels of cooperation even at low levels of forgiveness. The embedded indirect reciprocity strategy needs a higher probability to offer a second chance to sustain the same level of cooperation. Second, the embedded indirect reciprocity strategy (CR) is unable to sustain cooperation when strategies evolve fast while, in some cases, the blind strategy could overcome the pressure provided by the high speed of evolution. The reason for these differences can be attributed to the different set of information accessible to the two indirectly reciprocal strategies. Indeed, the blind strategy can access a broader network to condition its behaviour and thus can adapt its behaviour fast and acquire more information through forgiveness. The embedded strategy relies only on a smaller set of common partners, thus the acquired information is more local and less diverse, making it more difficult to discern pure defectors and other conditional players.

A limitation of our study is that we studied relatively small networks, in which indirect reciprocity might be less important than direct forms of reciprocity. Furthermore, small populations might also be prone to drift-like behavior resulting in local equilibria that would not be feasible in larger populations. However, we selected the relatively small network size because this corresponds better to the natural size of human groups, in the evolutionary past in particular. Our results are shown to be robust to even strong changes in the mechanisms of network dynamics underlying the interactions among the agents.

#### REFERENCES

- Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.
- Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.
- Boyd, R. and Richerson, P. J. (1989). The evolution of indirect reciprocity. *Social Networks*, 11(3):213–236.
- Dunbar, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of human evolution*, 22(6):469–493.
- Grafen, A. (1984). Natural selection, kin selection and group selection. *Behavioural ecology: An evolutionary approach*, 2:62–84.
- Gunnthorsdottir, A. and Rapoport, A. (2006). Embedding social dilemmas in intergroup competition reduces free-riding. *Organizational Behavior and Human Decision Processes*, 101(2):184–199.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52.
- Hauert, C. and Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snow-drift game. *Nature*, 428(6983):643–646.
- Hoffmann, R. (2000). Twenty years on: The evolution of cooperation revisited. *Journal of Artificial Societies and Social Simulation*, 3(2):1390–1396.
- Lieberman, E., Hauert, C., and Nowak, M. A. (2005). Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316.
- Nowak, M. and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, pages 1291–1298.
- Nowak, M. A. and Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685):573–577.
- Puurttinen, M. and Mappes, T. (2009). Between-group competition and human cooperation. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1655):355–360.
- Righi, S. and Takács, K. (2014). Emotional strategies as catalysts for cooperation in signed networks. *Advances in Complex Systems*, 17(02):1450011.
- Sachs, J. L., Mueller, U. G., Wilcox, T. P., and Bull, J. J. (2004). The evolution of cooperation. *The Quarterly Review of Biology*, 79(2):135–160.
- Santos, F. C., Pacheco, J. M., and Lenaerts, T. (2006). Cooperation prevails when individuals adjust their social ties. *PLoS Computational Biology*, 2(10):e140.
- Sigmund, K. (2012). Moral assessment in indirect reciprocity. *Journal of theoretical biology*, 299:25–30.
- Smith, E. A. (2010). Communication and collective action: language and the evolution of human cooperation. *Evolution and Human Behavior*, 31(4):231–245.
- Sommerfeld, R. D., Krambeck, H.-J., Semmann, D., and Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the national academy of sciences*, 104(44):17435–17440.
- Wedekind, C. and Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, 288(5467):850–852.
- West, S. A., Griffin, A. S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of evolutionary biology*, 20(2):415–432.
- Whitaker, R. M., Colombo, G. B., Allen, S. M., and Dunbar, R. I. M. (2016). A dominant social comparison heuristic unites alternative mechanisms for the evolution of indirect reciprocity. *Scientific Reports*, 6:31459.
- Wu, J., Balliet, D., and van Lange, P. A. M. (2015). When does gossip promote generosity? indirect reciprocity under the shadow of the future. *Social Psychological and Personality Science*, 6(8):923–930.



**SIMONE RIGHI** is Assistant Professor in Agricultural Economics at the University of Bologna and Researcher at the “Lendület” Research Center for Education and Network Studies of the Hungarian Academy of Sciences.

**KAROLY TAKÁCS** is Researcher and Director of the “Lendület” Research Center for Education and Network Studies of the Hungarian Academy of Sciences