

Reputation, Optimal Memory and the Evolution of Cooperation in Social Networks*

Gergely Horváth[†] Jaromír Kovářik[‡]
University of Alicante Basque Country University

Friederike Mengel[§]
Maastricht University

May 9, 2009

Abstract

In this paper we study a dynamic process where agents in a network interact in a Prisoner's Dilemma. The probability for two agents to interact decreases with their distance in the network. The network not only mediates interactions, but also information: Agents learn from their own experience and that of their direct neighbors about the history of any given player. Each agent can only memorize the last h periods. Action choices are rational given the preference type of the chooser and the reputation of the match. Evolution selects among preference types. We establish almost sure convergence of the process and characterize absorbing states. We then show - via simulations - that the probability of reaching a cooperative state does *not* relate monotonically to the memory constraint h . In fact the optimal memory level from a population viewpoint can be interior. We show that this is due to "stigma" i.e. to the fact that agents - even after changing their type are stuck for some time with their "old" reputation. Without "stigma" the probability of reaching a cooperative state increases monotonically with higher memory.

JEL-classification: C70, C72, C73.

Keywords: Evolution, Reputation, Optimal Memory, Cooperation.

*We are grateful to Jonas Hedlund, Hubert János Kiss, Marco van der Leij, Nick Vriend and seminar participants in IMEBE 2009 for helpful comments and suggestions. Jaromír Kovářik acknowledges Spanish Ministry of Science and Innovation (SEJ2006-06309/ECON) and the Government of the Basque Country (IT-223-07) for financial support.

[†]Corresponding Author. Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, 03080 Alicante, Spain. *E-mail*: gergely@merlin.fae.ua.es

[‡]Dpto. Fundamentos Analisis Económico I, Universidad del País Vasco, Av Lehendakari Aguirre 83, 48015 Bilbao, Spain. *E-mail*: jaromirkovarik@merlin.fae.ua.es.

[§]Department of Economics (AE1), Faculty of Economics and Business Administration, Maastricht University, PO Box 616, 6200MD Maastricht, Netherlands. *E-mail*: F.Mengel@algec.unimaas.nl

1 Introduction

Motivation

Human cooperation differs from cooperation observed in other species due to the use of mechanisms such as explicit reputation building or indirect reciprocity (reciprocity mediated through reputation..). It has been argued in the literature that this is due to the cognitive limitations of non-human species, preventing them to engage in more cognitively complex cooperation-enhancing mechanisms (Fehr and Gächter, 2002; Stevens and Hauser, 2004, 2005). Infact, in Biology it is widely recognized that human capacities, such as memory and language might be crucial for the evolution of cooperation (Trivers 1971, Nowak and Sigmund 1998). Existing literature in Evolutionary Game Theory, though, has largely ignored these issues.

In this paper we focus on one such human capacity and explore the role of memory for the emergence of cooperation. Our model is set in the tradition of the indirect evolutionary approach (See e.g. Bester and Gueth (1998)). In these models evolution selects among preference types, not strategies. In our model there are three types: altruists, defectors and conditional cooperators. Agents choose their actions as a best response to their beliefs about their current opponent's play.

To analyze the relevance of limited memory for prosocial behavior, we introduce two natural mechanisms in our model: First, we assume that agent i 's beliefs about j 's behavior are determined through the reputation j enjoys with i . Second, both interactions of agents as well as the spread of information are mediated through a social network. Reputation is created both directly, through repeated interactions of the same individuals, and indirectly, through information from neighbors in the social network. Agents have limited memory and remember only their last h interactions.

Finally - unlike in standard models - agents do not only interact with their first order neighbors in the network, but have positive probability to interact with anyone in the same connected component of the network. The probability to interact with a particular agent decreases in the "geodesic distance" of this agent. This allows us to capture situations where people are simultaneously engaged in various relationships. This certainly realistic feature of the model interacts with memory constraints in obvious and important ways. Both longer time intervals between potential cooperative encounters and having to recall the past play of multiple opponents are challenges for reputation-based mechanisms.¹

Our main results are the following. We establish almost sure convergence and show that absorbing states are characterized either by full defection or full cooperation. Furthermore, we explore the role of memory size, the importance

¹See Stevens and Hauser (2004) for a discussion along these lines and Milinski and Wedekind (1998) for experimental evidence on the effects of memory constraints on cooperative behavior.

of the network for matching as well as the role of "stigma" through simulations. "Stigma" refers to situations in which an agent after changing her type is still burdened by her "old" reputation. This case can reflect cultural evolution among humans and is certainly the correct model if changes of type are not observable. The "No-Stigma" case refers to the case where an agent after changing her type enjoys a "clean" reputation. This case can reflect biological (genetic) evolution or cultural evolution if changes of type are observable.² Our results show that in the "Stigma" Case there exists an interior memory size maximizing the probability to observe cooperative outcomes in the long run. The more important the network is for matching, the more cooperation is observed at the optimal memory size. In the "No Stigma" Case, on the other hand, a larger memory always leads to better chances for cooperation.

These results add an interesting new viewpoint to the discussion of why evolution did not endow us with unlimited brain capacity or, more loosely speaking, did not make us "infinitely" smart. Not only is it the case that larger brain or memory capacity involves a direct cost, our results also show that it can be optimal (at least from a population viewpoint) to be endowed with finite memory in strategic situations. In line with the group selection literature in Biology limited memory in our model can emerge from the conflict of two populations endowed with different memory capacity in the brain. Since human memory capacity is in reality limited, our results also suggest that stigma may play an important role in the evolution of human cognitive capacities.

The paper is organized as follows. Next we will discuss related literature. In Section 2 we will then present the model, in Section 3 explain our simulation strategy and in Section 4 present our main results. Section 5 concludes. Proofs and Regression tables are relegated to an Appendix.

Related Literature

Cooperative behavior based on reputation has been studied for decades especially in repeated bilateral interactions. So called tit-for-tat players cooperate if and only if their partner has cooperated in all past interactions. Studies include Trivers (1971), Rubinstein (1979), Axelrod and Hamilton (1981), Kreps et al. (1982) or Fudenberg and Maskin (1986). Kreps et al. (1982) for examples show that if for any agent there is a small probability to interact with such a tit-for-tat agent cooperation can be sustained even in the finitely repeated Prisoner's Dilemma.

However, cooperative behavior is also observed in one-shot interactions with unknown individuals. As a result, literature, which explores the role of reputation in non-regular encounters, has emerged. This led to the concept of indirect reciprocity (Nowak and Sigmund, 1998; Leimer and Hammerstein, 2001; Wedekind and Milinski, 2000; Seinen and Schram, 2004; Milinski et al., 2001): the reputation an agent i has in the eyes of an agent j also considers how i behaves with other individuals $k \neq j, i$. Under this concept, cooperation is fully

²Note, though, that with the explicit reputation building our model is essentially a model of *human* evolution.

constructed on reputation. The drawback of these studies is that the history of each player's actions is *public* knowledge. Hence, this mechanism needs an institution, which will disseminate information. An exception are Nowak and Sigmund (1998) who simulate a situation, where just a subset of the population learns the play in past bilateral encounters and show that public knowledge is not necessary for the cooperation to establish. In their model, the "observers" are chosen randomly from the population.³

In our paper there is a natural mechanism to disseminate information. Agents are located on a network and obtain information either through their own experience or through their direct neighbors in the network. All agents have limited memory, though, implying that reputation will always remain incomplete. Limited memory has been studied first by Miller (1956). He found that humans typically can remember seven objects of the same kind. Since then numerous studies have repeatedly reconfirmed this finding in a great variety of contexts, for a great variety of recalling objects, and across different age categories. Cowan (2001) presents a recent exhaustive review of this topic.

The emergence of cooperation in exogenous social networks has been studied by Eshel, Samuelson and Shaked (1998) or Fosco and Mengel (2008) among several others.⁴ One of the more closely related studies is Cortens and Cook (2008), who also model reputation building through the network. They study a dynamic network and show that in this setting reputation does not always enhance cooperation. Fosco and Mengel (2008) obtain a result with a similar flavor, showing that more information (not reputation) can hurt cooperation if agents imitate successful actions in an endogenous network.

Direct reputation-building within simple network architectures, such as circles and lattices has been studied by Boyd and Richerson (1989) or Nakamaru et al. (1997,1998). More recently, Ohtsuki et al. (2006) and Ohtsuki and Nowak (2007) have proceeded analytical results for regular graphs and simulation results for random and scale-free networks. The general conclusion of this literature is that, in most cases, the effect of the bilateral reputation is enhanced by the underlying network structure compared to the random matching case. See also Cohen et al. (2001).

Mohtashemi and Mui (2003) explicitly focus on the effect of social information that travels through network on cooperation. More precisely, direct

³A natural implication of the above assumption is that agents might have *no* information about their new opponent. The tit-for-tat strategy prescribes cooperation in such a case. Pollock and Dugatin (1992) propose an observer-tit-for-tat strategy that behaves as a tit-for-tat when nothing is known about the partner, but defects when the partner is known to have defected in the past. Nowak and Sigmund (1998) and Mohtashemi and Mui (2003) set this initial "reputation" fixed and most discriminative. However, it has been observed in the experimental literature that subjects' actions depend on what happens, on average, in the population. For instance, Fischbacher et al. (2001) and Kurzban and Houser (2005) report that conditional cooperators, the only types in our model that make use of information, contribute according to what the others do on average. Other literature, moreover, argues that humans have an innate capacity to detect cheaters (Cosmides and Tooby, 1992). Insofar, we assume that the average rate of cooperation is publicly known and when ignorant about the past play of opponents agents base their decision of the average rate.

⁴Further literature can be found among the references of these articles.

links mutually share information in their model, spreading reputation of past encounters across neighbors. They rule out repeated interactions and show that non-direct assesment of reputation itself can stabilize cooperative behavior. However, they model a growing network, in which everybody ends up knowing everybody, which is crucial for the survival of cooperation in their set-up.⁵

Finally, in an experimental study, Cassar (2007) contrasts cooperation levels across three networks: local, small-world, and random. She reports that the initial frequency of cooperation lies around 50% in all networks and decreases over time, suggesting that the social network structure cannot sustain cooperation on its own. This decrease is actually more pronounced in the small-world network and similar in the other two network types.

2 The Model

There is a finite number of agents interacting in a bilateral Prisoner's Dilemma through a given network. Agents are more likely to interact with others that are "closer" to them in the network. (Subsection 2.2). There are different types having different preferences over the outcomes in the Prisoner's Dilemma. (Subsection 2.1). Given their preferences agents best respond to their beliefs about their opponent's choice. Beliefs are determined explicitly through the reputation this opponent enjoys with the agent in question. (Subsection 2.3). Ultimately we assume that evolution selects among preference types. (Subsection 2.4).

2.1 Game and Payoff Types

The payoffs of an agent of type τ when choosing action a_i and his opponent chooses a_j are given by

$a_i \backslash a_j$	C	D	
C	a	0	(1)
D	$1 - w_\tau$	$d - w_\tau$	

with $1 > a > d > 0$. We also assume that $d > 1 - a$.⁶ If $w_\tau = 0$ this matrix represents a (standard) Prisoner's Dilemma. Different values of w_τ can induce different types of behavior. We assume that there are three types in the population. The type space is $T = \{D, CC, A\}$.

- D : Defectors ($w_D < 1 - a$) whose optimal strategy is always defect.
- CC : Conditional cooperators ($w_{CC} \in [1 - a, d]$). Conditional Cooperators cooperate whenever their beliefs about the probability that their next

⁵They themselves remark that without the network growth cooperators are easily defeated by defectors (Mohtashemi and Mui, 2003, p. 527).

⁶This simplification serves to determine the behavior of conditional cooperators in a natural way. See Mengel (2008).

round match cooperates σ is larger than the critical value $\hat{\sigma}$:

$$\hat{\sigma} = \frac{d - w_{CC}}{a + d - 1}.$$

- A: Altruists ($w_A > d$) who will always cooperate.

2.2 Interaction Structure

Agents are organized on an exogenous undirected network. At each period one agent is chosen randomly (with uniform probability) to play the Prisoner's Dilemma. This player is matched with somebody else from the network to play the game. The probability to interact with any player is proportional to the shortest "geodesic distance" that separates the two players.⁷ Denote the shortest distance between i and j by d_{ij} (obviously $d_{ij} = d_{ji}$). Then the probability for i to interact with j given that i has been chosen as rowplayer is the following:

$$P_{ij}(d_{ij}) = \begin{cases} \frac{e^{-\alpha d_{ij}}}{\sum_{k \neq i} e^{-\alpha d_{ik}}} & \text{if } d_{ij} > 0 \\ 0 & \text{if } d_{ij} = 0 \text{ or } d_{ij} = \infty \end{cases}$$

Note that with this function $P_{ij}(d_{ij}) \neq P_{ji}(d_{ji})$, more central players in the network are matched more frequently.⁸ Evolutionary fitness, though, will depend only on the expected payoff per interaction, ensuring that "centrality" does not *per se* lead to higher fitness. Also note the following features of the model. The higher is α the more likely it is to meet neighbors and the more unlikely to meet distant individuals. α thus parametrizes the effect of the network for matching. If $\alpha = 0$ matching is random with uniform probability - so we have eliminated the network effect. On the other hand if $\alpha \rightarrow \infty$, matching takes place exclusively among direct neighbors. Finally, note that if $d_{ij} = \infty$ (agents in disconnected components) then $p(d_{ij}) = 0$.

When playing the game altruists (A) always choose cooperate and defectors (D) always choose defect. For conditional cooperators (CC), though, action choice depends on their beliefs on what their match will choose. These beliefs in turn depend on the reputation their match enjoys with them.

2.3 Reputation and beliefs

Agents have limited memory. They remember their last h interactions. They form beliefs about the behavior of the opponent player using their experience

⁷Geodesic distance between any two nodes is the minimum number of links that connects the two nodes.

⁸Given that i has been chosen as rowplayer, the probability that i is matched with somebody else adds up to 1: $\sum_{j \neq i} P_{ij}(d_{ij}) = \sum_{j \neq i} \frac{e^{-\alpha d_{ij}}}{\sum_{k \neq i} e^{-\alpha d_{ik}}} = 1$.

from the last h periods and the information they can get from their direct neighbors.

Denote by $\beta_{ij}(h)$ the number of times that i interacted with j in the last h periods and $\gamma_{ij}(h)$ the number of times that j cooperated in such an interaction. If there was no interaction between i and j in the last h periods $\beta_{ij}(h) = \gamma_{ij}(h) = 0$.

There is no global "reputation" in the model. Some players will think badly about a given player and others good. Denote $N_i = \{k \in N \mid d_{ik} = 1 \wedge \beta_{kj}(h) > 0\}$ and $n_i(h) = \text{card}(N_i)$. This latter is the number of player i 's first-order neighbors in the network that have interacted with agent j in the last h periods. The reputation that player j has for player i (what i thinks about j) at time t is given by

$$r_{ij}^t = \begin{cases} \lambda \frac{\gamma_{ij}(h)}{\beta_{ij}(h)} + \frac{(1-\lambda)}{n_i(h)} \sum_{k \in N_i} \frac{\gamma_{kj}(h)}{\beta_{kj}(h)} & \text{if } \beta_{ij}(h) > 0 \wedge \\ & n_i(h) > 0 \\ \frac{\gamma_{ij}(h)}{\beta_{ij}(h)} & \text{if } \beta_{ij}(h) > 0 \wedge \\ & n_i(h) = 0 \\ \frac{1}{n_i(h)} \sum_{k \in N_i} \frac{\gamma_{kj}(h)}{\beta_{kj}(h)} & \text{if } \beta_{ij}(h) = 0 \wedge \\ & n_i(h) > 0 \end{cases} .$$

The reputation that j enjoys with i is thus a weighted average of the proportion of times that j cooperated with i during the last h periods and the number of times that j cooperated with i 's friends. The idea is that i has information from his neighbors about j (given that they interacted with j). He forms an opinion about j as a weighted average of his own experience with j and the experience of his friends. High λ means that he relies mostly on his own experience and low λ that he forms judgements based on the information from others.

If nothing is known about j (i.e. if $\beta_{ij}(h) + \sum_{k \in N_i} \beta_{kj}(h) = 0$) there is "no reputation". In this case the agents use the average rate of cooperation in period $t-1$ which is assumed to be always known to all players ($\bar{\sigma}^{t-1}$). In other words the default belief about others (when nothing is known about them) is the population average.

In sum, the belief that i has about the likelihood that j cooperates at t is given by

$$\sigma_{ij}^t = \begin{cases} r_{ij}^t & \text{if } \beta_{ij}(h) + \sum_{k \in N_i} \beta_{kj}(h) > 0 \\ \bar{\sigma}^{t-1} & \text{otherwise} \end{cases} .$$

Our set up relates the network structure to the information accumulation in a natural way. Given the locally biased matching process, if clustering is high or the network is small, it is very likely that the individual will know something about j ($\beta_{ij}(h) + \sum_{k \in N_i} \beta_{kj}(h)$ will be strictly positive). In this case agents are able to identify the type of the opponents and probably cooperation has good chances to survive. In the other extreme case the players are likely to know very little about the opponent and they form opinion by looking on the population average. Then the model is close to an incomplete information set up, in which the chances for cooperation should be rather bad.

2.4 Selection process

Ultimately, we are interested in which of the three types survive in an evolutionary model. Evolutionary success (fitness) is measured by per interaction material payoffs of the agents. Obviously fitness can depend only on material payoffs (not taking into account w_τ), as we want to know which preferences will be selected. As mentioned before we use per interaction payoffs such that more "central" agents or agents with a higher degree *do not per se enjoy* higher fitness.

More specifically the selection process is modeled as follows. Every N periods, i.e. once each agent had the chance to play the Prisoner's Dilemma exactly once on expectation, k agents are selected randomly for selection. For each of these agents one agent is randomly chosen from the whole population and their fitness is compared. If the randomly chosen agent has higher fitness, the agent adopts his type. This process could be interpreted as imitation learning, but also as genetic evolution.

We consider two alternative scenarios. In our benchmark scenario (Section 4.1) we assume "Stigma", meaning that an agent after switching her type still has the same reputation as before. Also the memory of the agent is not erased. This assumption is closer to a setting of cultural evolution, e.g. a process of imitation.

Then (in Section 5.2) we also consider the "No-Stigma Case" in which an agent's reputation is "cleaned" after she changes her type. This case could correspond to biological (genetic) evolution but also to cultural evolution if type changes are observable (through either changes in action choices or coproducts outside this model).

2.5 State Space and Convergence

Note that at each point in time t each player is entirely characterized by the vector of her reputations for others $r_i(t) := (r_{i1}^t, r_{i2}^t, \dots, r_{iN}^t)$ as well as those of her first-order neighbors in the network. Consequently the population state at time t is given by the $n \times (n + 2)$ matrix

$$\Xi(t) = (\tau(t) \parallel \pi(t) \parallel R(t)),$$

where $\tau(t) = (\tau_1(t), \dots, \tau_N(t))^T$ is the $n \times 1$ vector indicating the types of all players, $\pi(t)$ the $n \times 1$ vector indicating each agent's payoff from the last interaction and $R(t) = (r_1(t), \dots, r_N(t))^T$ is the $n \times n$ matrix indicating all the player's reputations for each other. Note that since $\Xi(t)$ is completely determined by $\Xi(t - 1)$ and the realizations of the random variables at t , the associated transition matrix describes a finite Markov chain on the state space $S := T \times \Pi \times \mathcal{R}$ (where \mathcal{R} is the set of all possible reputation matrices and Π is the payoff space). Denote the probability to reach state s' from state s by $q(s, s')$. We have the following definition.

Definition State s is absorbing $\iff q(s, s) = 1$.

Our first result shows that the stochastic process does indeed converge to one of these absorbing states.

Proposition 1 *If α is finite, the stochastic process described above converges almost surely to an absorbing state.*

Proof. Appendix. ■

This is an important result, since it enables us to be confident that our simulations should converge. In the next proposition we give a partial characterization of the absorbing states.

Proposition 2 *If α is finite, every absorbing state is characterized by either full defection or full cooperation.*

Proof. Appendix. ■

Proposition 2 shows that strong "norms" of behavior will tend to evolve. Either everyone will end up cooperating or everyone will end up defecting. This of course immediately implies that defectors and altruists cannot coexist in any absorbing state. (Since defectors always defect and altruists always cooperate). A different question is of course - which of these states are we likely to observe in the long run? This will naturally depend on the relative size of their basins of attraction. In the next sections we describe the results of simulations, that will show us which of these states are likely to emerge and how the basins of attractions of the different states vary with the parameters of our model. Particular importance will be attached to the memory size h of the agents.

3 Simulations

We simulate the model using small world network (Watts and Strogatz (1998)) with $N = 100$ agents.⁹ The small world network is generated from a one-dimensional lattice where the agents have links to their neighbors of up to ρ distance (ρ is the connection radius). After each link is rewired with some probability (θ): the link between agent i and j is destroyed and agent i is connected to a randomly chosen other agent who is not already connected to i .

The rewiring procedure has an important impact on network measures, such as the clustering coefficient and the average distance. The clustering coefficient of a network is the average of the clustering coefficient of the nodes. The latter is the fraction of pairs of neighbors of a given node that are themselves neighbors (see Vega-Redondo (2007) Ch. 2). In the case when the network is connected, the average distance is the average of the length of the shortest path between any two nodes, i.e. the average of the geodesic distance over the pairs of nodes. When the network is disconnected, the average distance is infinite according to the conventions.

⁹Simulations are programmed and carried out using the program RePast.

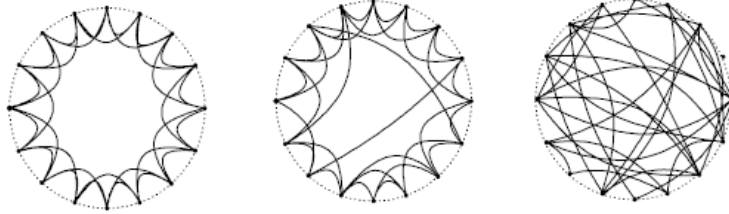


Figure 1: Network structures for $\rho = 2$ and different values of rewiring probability. Left: $\theta = 0$, middle: $\theta = \epsilon$, right: $\theta = 1$. Taken from Vega-Redondo (2007) p. 59.

If the rewiring probability θ is 0, we have a local network which is characterized by high clustering and high average distance and each node has 2ρ neighbors. Nodes form relatively isolated neighborhoods. As the rewiring probability θ takes a positive but small value, the average distance decreases dramatically: a few shortcuts connect the local neighborhoods which makes it easier to reach all nodes in the network from a given one. However, this doesn't change the clustering coefficient significantly. This type of small world phenomenon has been observed in many real social networks. Figure 1 illustrates the effect of rewiring.

When the rewiring probability is high, the local neighborhoods get dissolved, the clustering coefficient is small. The average distance takes the small value typical of the random networks. Due to the rewiring this network is characterized by a heterogenous degree distribution. Table 1 shows the mentioned measures for the simulated networks.

In our simulations, the connection radius of the ring network to be rewired takes the values 2 and 4, the rewiring probability changes between 0 and 1.

Initially each type has equal share in the population. Types are distributed randomly over the network. In the first period, the default belief is set to $2/3$ (the sum of the share of cooperators and of conditional cooperators).

At each period one agent is chosen randomly as row player and is matched with an other agent using her matching probabilities based on her network position. The locality of the matching α was fixed to be 2 (see the example in the Appendix about the effect of the magnitude of α). Agents play the Prisoner's Dilemma once. The game's payoffs are specified to $a = 0.8$, $d = 0.5$, $w_{CC} = 0.35$. Players of conditional cooperator type use the available information or the default belief to determine their action. In case when the player and her neighbors as well have information about the opponent, the two sources have equal weights in determining the player's beliefs ($\lambda = 0.5$). The row player's payoff and memory status (experience) is updated according to the interaction. The column player's status remains unchanged.

Connection radius = 2			
Rewiring Prob.	Clustering	Avg Distance	% connected graphs
0	0.5	13.009	100
0.01	0.475	8.808	100
0.05	0.384	5.219	100
1	0.037	3.488	17.2
Connection radius = 4			
Rewiring Prob.	Clustering	Avg Distance	% connected graphs
0	0.643	6.765	100
0.01	0.611	4.391	100
0.05	0.498	3.144	100
1	0.080	2.450	98.1

Table 1: Measures of the used networks. The data are averages from 10.000 observations of the network for each value of the network parameters. Average distance is computed using the subsample of connected graphs.

After N periods of interaction each agent had the chance to move once on expectation. Then $k = 20$ agents' type is updated according to the selection procedure where fitness is associated to previous period's payoff. The default belief is updated to the average number of cooperations in the last N periods. Agents forget the information about the h^{th} oldest interaction in their memory. After there are other N periods of playing the Prisoner's Dilemma and so on.

We simulated the system for 20000 periods (i.e. each agent has moved 20000 times) and we run the model 100 times for each parameter set. We looked at the number of cooperating agents in the final period and at the final type distribution. We were interested in analyzing the effect of memory (h) and of network structure on these outputs. Later we analyzed the sensitivity of the results to the changes of other parameter values (α , λ , k , w_{CC}). We also investigated the No Stigma case, when the type updating agents do not carry over their memory and reputation to the following periods.

Table 2 summarizes the parameters applied in the basic runs of the model.

4 Results

4.1 Memory and Network Structure

In our model the survival of cooperation crucially depends on the information accumulation process which is needed for the conditional cooperators to be able to differentiate the defectors and the altruists. The efficiency of this process depends on the following parameters: memory (h), rewiring probability (θ , hence clustering), connection radius (ρ), speed of selection (k) and the locality of the matching (α).

Parameter	Value
N	100
a	0.8
d	0.5
w_{CC}	0.35
λ	0.5
k	20
α	2
ρ	2, 4
h	[1,20]
θ	0, 0.01, 0.05, 1

Table 2: Parameters in the basic runs

First we focus on the "Stigma" case and analyse the effect of memory on the level of cooperation under different network structures represented by different values of connection radius and rewiring probability. The influence of other parameters is analyzed in the section of sensitivity analysis. Remember from Proposition 2 that all absorbing states are either characterized by full defection or by full cooperation. When we refer to the frequency of cooperative states we mean the frequency with which fully cooperative states were observed as an outcome of our simulations. Result 1 describes the case where agents have few connections.

Result 1 *When $\rho = 2$, there exists values of memory h_1 and $h_2 > h_1$ s.t. if $h < h_1$ a cooperative state is never reached. For $h \in [h_1, h_2)$ the frequency of cooperative states is increasing in h and remains at the maximum level for $\forall h \geq h_2$. At h_2 convergence to cooperative absorbing states is fastest for $\forall h \geq h_2$. Both h_1 and h_2 are increasing in the rewiring probability, while the frequency of cooperative states decreases in rewiring probability.*¹⁰

Figure 2 shows the frequency of cooperative states for different memory and rewiring probability values. For $\theta \leq 0.05$, memory should be at least 3 in order to have cooperative outcomes. For $\theta = 0$ and $\theta = 0.01$, the maximal value of cooperation is around 60-65 which is reached at memory 6-7. As the rewiring probability increases, h_1 increases being 7 for $\theta = 1$. h_2 is 8-9 for $\theta = 0.05$ by which cooperation reaches the level of 45-50. For $\theta = 1$, the maximal value of cooperation is reached around $h_2 = 17$ taking the value of 35-40.

In each run the model converges to one of the two extrem states: full cooperation or zero cooperation. The average time to convergence for different memory and rewiring probability values is shown in Figure 3. Since the time of convergence monotonically increases in memory, h_2 represents the memory value by which the maximum possible cooperation is reached within the shortest time period. This would be the socially desired value of memory length.

¹⁰Obviously, h_1 and h_2 are functions of all the model parameters.

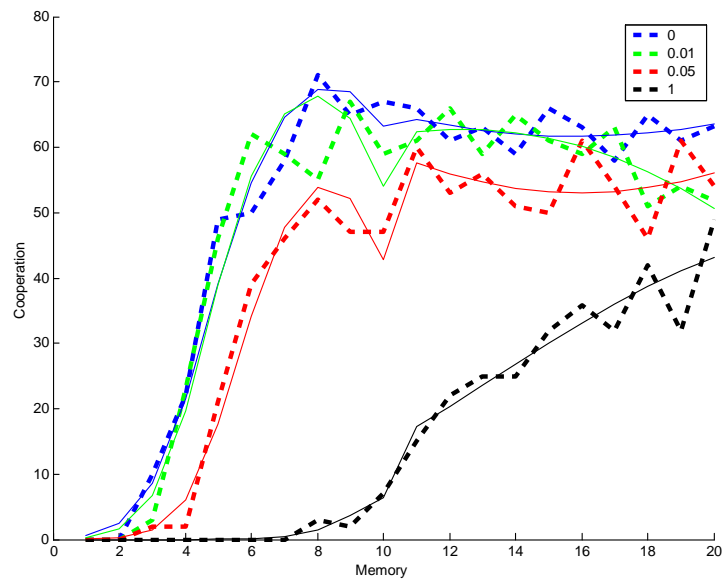


Figure 2: Frequency of cooperative outcomes as a function of memory for different rewiring probabilities when $\rho = 2$. Dashed line: data, solid line: fitted values from the econometric model (see below).

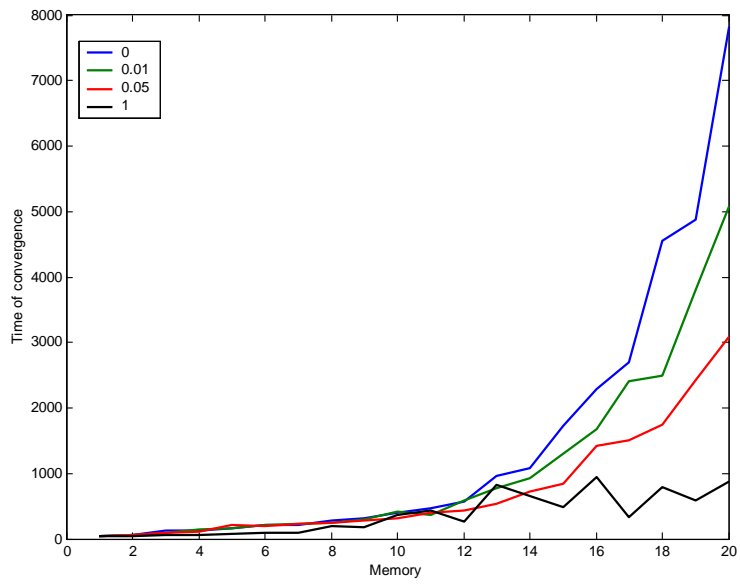


Figure 3: Average time to convergence as a function of memory for different rewiring probabilities (average of 100 runs per parameter setting).

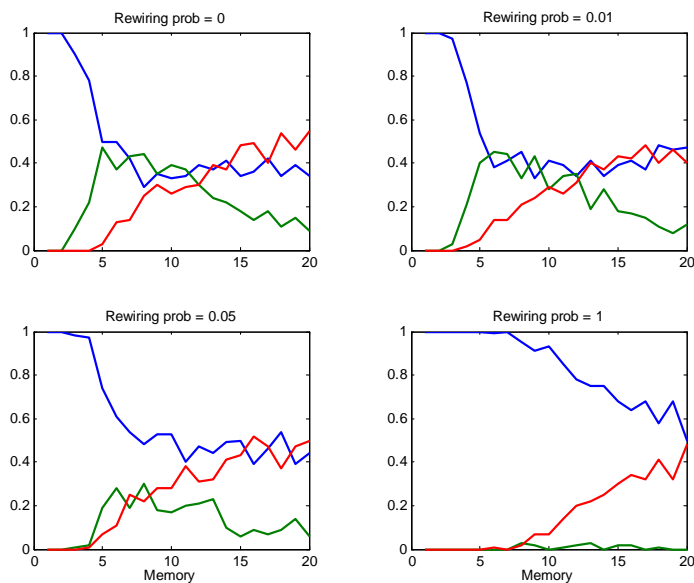


Figure 4: Shares of the different scenarios among the 100 runs per parameter setting as a function of memory for different rewiring probabilities if $\rho = 2$. Blue line: zero cooperation. Green line: full cooperation. Red line: CC survive.

Figure 4 shows the final type distribution classified into 3 cases: 1. "Zero cooperation" (blue line) when conditional cooperators and defectors survive, 2. "Total cooperation" (green line) when conditional cooperators and altruists survive, 3. "CC survive" (red line) when just conditional cooperators survive. The graph shows how the shares of these scenarios change as memory and rewiring probability takes different values. In the first case the model always reaches the state when no agent cooperates while in the second and third case it converges to full cooperation. By low values of memory the altruists die out always. As memory increases defectors extinct more frequently but as memory gets very large the share of the case when just the conditional cooperators survive increases steadily. In this later region the altruists cannot survive again. When the rewiring probability θ is very large altruists almost never survive.

The intuition behind Result 1 can be understood from the mechanisms of the model. When memory is very low ($h < h_1$), conditional cooperators cannot learn the type of the opponent players they are most frequently matched with. As visual observation of the running model indicates, at the beginning of the simulations defectors exploit altruists, this yields the highest payoff for them, so they can increase their share in the population. As the share of altruists

decreases they are mostly matched with defectors and conditional cooperators. These later eventually defect against them because of the lack of information¹¹. This leads to the extinction of altruists and to zero cooperation since defectors and conditional cooperators will play Defect against each other.

As memory increases ($h > h_1$) CC agents can prevent the extinction of altruists. After the initial expansion, defectors begin to earn low payoff as they are mostly matched with defectors and conditional cooperators. Due to more information provided by the memory, CC type agents now are able to differentiate the A and D type agents, they defect against defectors and cooperate with cooperators. This gives high payoff for altruists and low payoff for defectors leading to the expansion of altruists and the extinction of defectors. The share of "Total cooperation" case soar on Figure 2 and the rate of cooperation increases on Figure 1.

The length of memory necessary to the rise of cooperation depends on the network structure represented by the rewiring probability. When the rewiring probability is low, clustering is high in the network, shorter memory is sufficient to the increase of cooperation. Clustering and the locally biased matching rule imply that the information can accumulate locally: matching happens most likely with direct neighbors and by clustering the agent can monitor most of the actions of their neighbors even though she is matched just with one of them in each round. This is because she can gain information from their neighbors about their other neighbors who are mostly matched with each other as well¹². In this way, although in our model the reputation is not objective, in the sense that the perception of a person by the agents can be different, in a local neighborhood the reputation of a person becomes a shared opinion of everybody.

The lack of clustering cannot be offset by longer memory. When rewiring probability is large ($\theta = 1$), for $h > h_2$ the average cooperation is still lower than in the case when θ is low. For a given value of memory, cooperation always decreases in rewiring probability.

As shown below, the information accumulation is also supported by the fact that by $\rho = 2$ the agents have just a few neighbors they are frequently matched with and they have to learn about.¹³

The finding that clustering enhances cooperation in small world networks is supported by previous studies. Masuda and Aihara (2003) show that the rate of cooperation is higher by low rewiring probability when only defectors and altruists are in the population and agents imitate each other using a local payoff based mechanism. Here cooperators are able to form a massive cluster which increases gradually by the imitation process. The experimental paper of Cassar (2007) shows that people tend to cooperate more in a local network

¹¹As the altruists' share and the cooperation decrease, the conditional cooperators defect when they don't have information about the opponent because they use the average number of cooperations as default belief.

¹²Agents cannot learn the action of their neighbor from the neighbor herself as agents don't store their own action in their memory.

¹³In the lattice network to be rewired the number of neighbors for each agent is equal to 2ρ . If the rewiring probability is low, in the small world network the degree distribution is concentrated around this value.

(when $\theta = 0$) than in a small world network. However, the effect of rewiring seems to be non-monotone in reality as on the random network (when $\theta = 1$) individuals cooperate more than on small world network.

As memory gets very long ($h > h_2$), altruists die out more frequently, the cooperation is sustained by the CC type agents who are more often the only survivors. However, with only conditional cooperators the system converges much slower to the full cooperation state, and the time of convergence increases in memory. As visual observation of the running model indicates, in these runs typically first the altruists die out. This can be explained by the interaction between the length of memory and the selection process: as the selection changes the type of the agents, their "previous life's" decisions are still remembered which hinders the conditional cooperators in deciding the right action against them (see in contrast the "No Stigma" case).

At the beginning, defectors increase their share by exploiting altruists, this has three effects: 1. later defectors are matched with each other and begin to earn low payoffs, 2. altruists are matched frequently with defectors, 3. since the share of altruists decreases steadily, conditional cooperators are able to identify them (notice that the described effect of selection process is present just if there is inflow into the type's subpopulation). 1 and 3 imply that there will be some defectors turning to altruists. But then the conditional cooperators still are going to defect against them. This, together with 2, results in the extinction of altruists even though the memory is long.

After the extinction of altruists, with time the defectors become identifiable. Conditional cooperators will defect against them and cooperate with those CC type agents who have good reputation. Hence, the share of defectors is going to decrease which leads to the extinction of defectors.

The number of cooperating agents remains low because conditional cooperators have gained bad reputation by interacting with defectors and this later is remembered for a longer time by the long memory (and implies further defections). The system of CC type agents converges to full cooperation by a slow process where the time of convergence is increasing in memory. In the end the agents gain good reputation by cooperating with those CC type agents who had good reputation, cooperation diffuses on the network. In the case when altruists are also present this process is faster since CC agents can increase their reputation by cooperating with them. So, the time to convergence is smaller in this case.

Result 2 describes how things change when agents have many connections.

Result 2 *When $\rho = 4$, there exist h_1 and $h^* > h_1$ s.t. if $h < h_1$ a cooperative state is never reached. If $h \in [h_1, h^*]$ frequency of cooperative states is increasing in h and for $h > h^*$ it is monotonically decreasing. Both h_1 and h^* are increasing in the rewiring probability, while the frequency of cooperation decreases in the rewiring probability.*

In general, for every value of memory h and rewiring probability θ , the cooperation is decreasing in connection radius r . This is because agents have

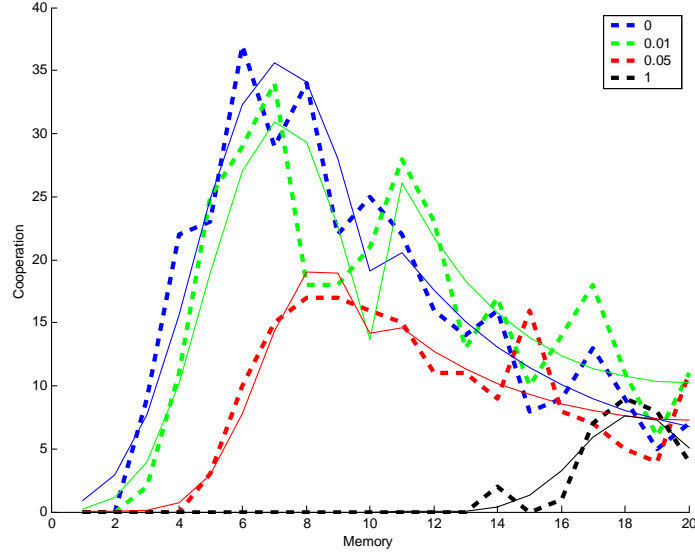


Figure 5: Frequency of cooperative outcomes as a function of memory for different values of rewiring probability when $\rho = 4$. Dashed line: data, solid line: fitted values from the econometric model (see below).

more neighbors to learn about. However, the main difference between the $\rho = 2$ and the $\rho = 4$ case is that in the latter there is a cutoff value of memory after which cooperation decreases. This value is increasing in θ . This can be seen on Figure 5.

Figure 6 shows the classification of the runs according to the survival of types. We can observe that with $\rho = 4$ the share of the case when just the conditional cooperators survive doesn't increase with memory. Instead, the defectors are able to survive by higher memory values as they do by very low memory.

For low values of memory ($h < h_1$) the model always converges to zero cooperation. For higher memory values ($h_1 < h < h^*$) altruists survive more frequently, the average cooperation increases as in the case of $\rho = 2$. However, this rise is lower in this case because conditional cooperators are able to differentiate defectors and altruists less frequently due to the increased number of neighbors they are matched with most often.¹⁴ As memory increases further ($h > h^*$), altruists die out often in the same way as in the previous case but now cooperation begins to decrease because defectors are able to survive as well. The process converges to zero cooperation before the defectors die out because

¹⁴Each agent has around 2ρ neighbors. If α is such that there is a significant probability to be matched with second order neighbors, agent should have information about $4\rho^2$ neighbors.

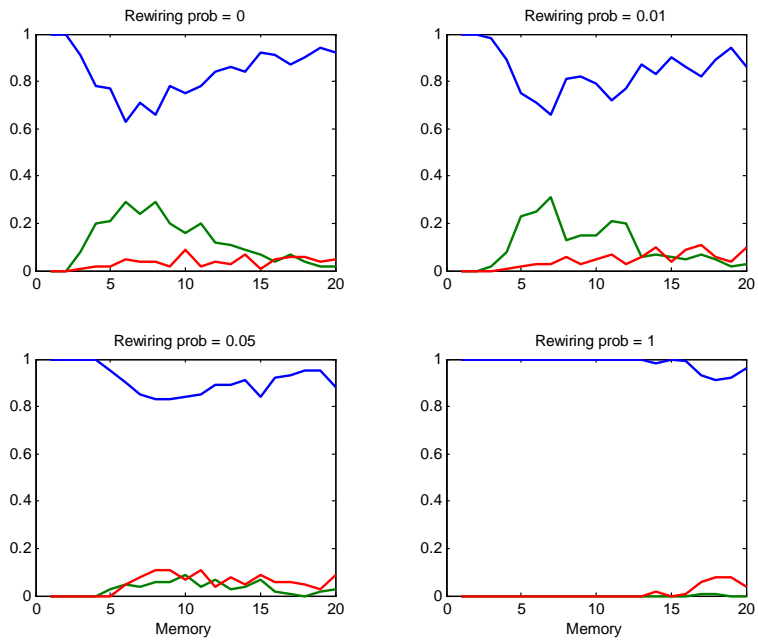


Figure 6: Shares of the different scenarios among the 100 runs per parameter setting as a function of memory for different rewiring probabilities for $\rho = 4$. Blue line: zero cooperation. Green line: total cooperation. Red line: CC survive.

due to the lack of information CC agents eventually defect against each other (at this point the average cooperation is very low, so in case of no information agents play defect). Cooperating CC type agents cannot diffuse cooperation because they are not recognized by others.

To demonstrate the difference between the two cases we regressed the level of cooperation on quadratic functions of memory using logit model given that the cooperation converges to 0 or 100 in each run of the model. We fitted two quadratic functions in memory with a breakpoint at $h = 10$. The regression tables are relegated to the Appendix (Table 4 and 5), the fitted values are shown on Figure 2 and 5. We used these models to test whether the marginal effect of memory is zero for $h > 10$ (i.e. $H_0: 2\beta_2h + \beta_1 = 0$ versus $H_1: 2\beta_2h + \beta_1 \neq 0$)¹⁵. The results are shown in Table 3 for different values of memory, connection radius and rewiring probability.

	$\theta = 0$		$\theta = 0.01$		$\theta = 0.05$		$\theta = 1$	
	$\rho = 2$	$\rho = 4$	$\rho = 2$	$\rho = 4$	$\rho = 2$	$\rho = 4$	$\rho = 2$	$\rho = 4$
h=10								
Test stat.	0.3	2.07	0.17	4.5*	0.88	1.21	3.88*	7.11*
P-value	0.5864	0.15	0.6824	0.0326	0.3482	0.2708	0.049	0.0078
h=13								
Test stat.	0.29	6.25*	0.04	9.73*	0.87	2.75	9.59*	7.78*
P-value	0.591	0.0124	0.8372	0.0018	0.35	0.098	0.002	0.0054
h=16								
Test stat.	0.00	10.36*	6.63*	9.43*	0.00	3.38	28.32*	9.75*
P-value	0.96	0.0014	0.01	0.002	0.969	0.0658	0.00	0.0018
h=20								
Test stat.	0.21	0.37	2.49	0.00	0.6	0.01	0.78	2.45*
P-value	0.646	0.542	0.1146	0.971	0.4386	0.9266	0.3774	0.0118

* denotes rejection of the null at 5% significance level.

Table 3: Test for the effect of memory on cooperation for $h > 10$

The tests show that in case of $\rho = 2$ we cannot reject the null that memory has no effect on cooperation when it is large while in case of $\rho = 4$, we do can. This reinforces the results what we have seen on the graphs. Notice that at $h = 20$ the level of cooperation becomes constant in memory in both cases, but at $\rho = 4$ it goes to a low value. When $\theta = 1$, for large values of memory the cooperation just begins to be positive, this explains why the tests reject the no-effect hypothesis in this case.

In sum, we have seen that longer memory is socially desirable up to a point but very long memory is harmful. In the case when agents have few connections,

¹⁵We used two-sided Wald-test.

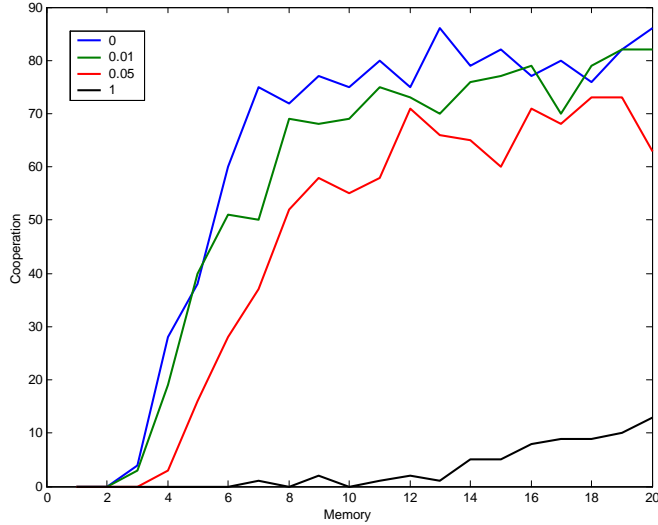


Figure 7: Average cooperation as a function of memory for different rewiring probability values when $\rho = 2$ in the "No-Stigma Case".

by longer memory the society has to wait too much for the high level of cooperation. However, when agents have many neighbors the too long memory results in the realization of defection equilibrium. In addition, by very long memory the altruists are never able to survive.

Regarding network structure, the high average degree is unfavorable for cooperation while the clustering helps the information accumulation and therefore the emergence of cooperation.

4.2 The "No Stigma" Case

As indicated above some of our results depends on the specified cultural evolutionary process, namely, that agents carry over their reputation after changing their type. In this section we analyze the model for a different mechanism when agents start their new life unexperienced (clean memory) and being unknown (clean reputation) for other players which corresponds to a biological situation. This assumption obviously improves the information accumulation process since in the case of long memory agents don't store false information about other agents who have changed their type in the meanwhile.

Figure 7 shows the average cooperation from 100 runs of the model for connection radius $\rho = 2$ and applying the "biological" selection process.

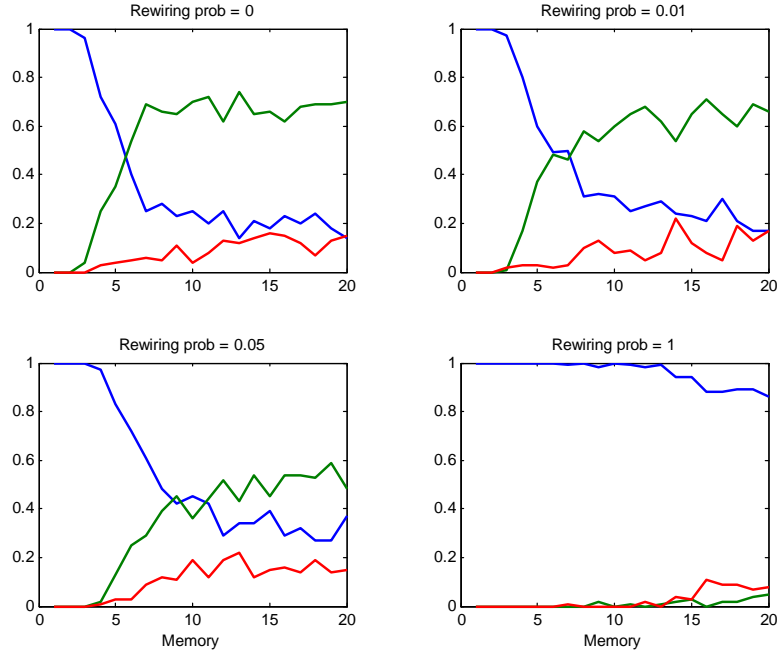


Figure 8: Shares of the different scenarios among the 100 runs per parameter setting as a function of memory for different rewiring probabilities for $\rho = 2$ in the "No-Stigma Case". Blue line: zero coepration. Green line: full cooperation. Red line: CC survive.

Comparing Figure 2 and Figure 7 reveals that at least for smaller values of rewiring probability the average cooperation is higher in this case than with cultural evolutionary process. This is in accordance with Axelrod's suggestion (Axelrod (1984)) that forgiveness is in favor of cooperation: to punish early defection for too long time easily traps the individuals in everlasting mutual retaliation. Figure 8 shows the final type distribution classified into 3 cases as above.

The main difference with cultural evolution is that here as memory increases the altruists are able to survive more frequently for lower values of θ . This results in the increase of cooperation for these parameter values. The biological evolution improves the conditional cooperators' possibility to differentiate between defectors and altruists since now they don't confuse the agents' actual type with their previous one. However, for large θ almost always the defectors and the conditional cooperators survive, the cooperation stays at very low level. This is because clustering is also needed for the learning of CC type agents.

When the connection radius is set to 4 and we apply biological evolution,

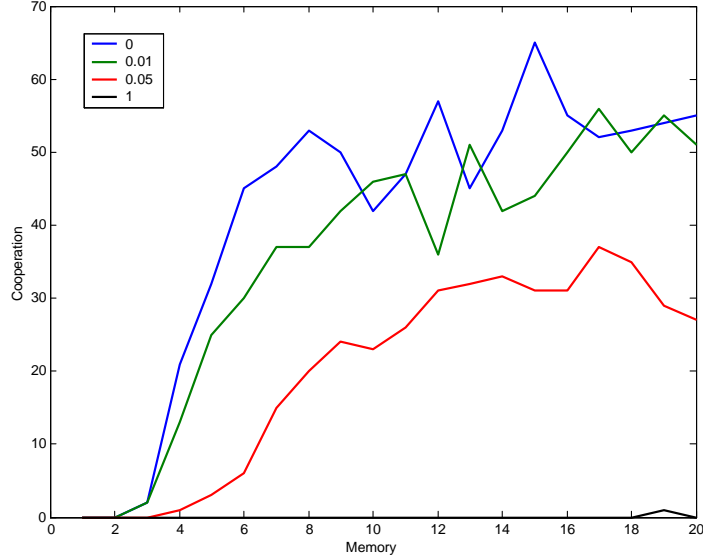


Figure 9: Average cooperation as function of memory for different rewiring probabilities when $\rho = 4$ in the "No-Stigma Case".

there doesn't exist h^* anymore: cooperation steadily increases in memory at least when clustering is not very low in the network. This is again because altruists survive and defectors die out more frequently as memory gets higher. This can be observed on Figure 9 and 10.

4.3 Sensitivity analysis

In this section we investigate the sensitivity of the results discussed before to the parameters that were fixed in the previous analysis: the locality of matching (α), the number of imitating agents (k), the weight of own and neighbors' information (λ) and the preference bias of CC type for cooperation (w_{CC}). We held every parameter fixed at the value indicated in the section of simulation but one which is changed on a grid. We consider the basic model with "No Stigma" selection process. The corresponding graphs are reported in the Appendix.

The locality of matching is an important parameter in the maintenance of cooperation. As α gets higher agents are matched most frequently with agents from the neighborhood which limits the number of possible opponents and facilitates the learning about their types. Therefore, as it can be seen on Figure 11 and 15, for a given value of h and θ the cooperation increases in α . The value of memory necessary to have some cooperation rises with α , however the overall pattern of cooperation as a function of memory is robust to changes

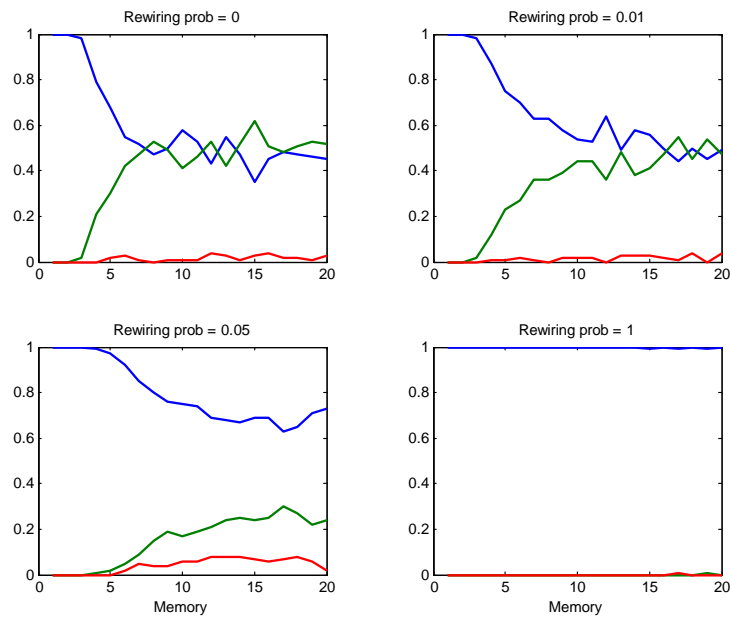


Figure 10: Shares of the different scenarios among the 100 runs per parameter setting as a function of memory for different rewiring probabilities for $\rho = 4$ in the "No-Stigma Case". Blue line: zero cooperation. Green line: full cooperation. Red line: CC survive.

in the locality of matching for both $\rho = 2$ and $\rho = 4$.

Figure 12 and 16 show the frequency of cooperative outcomes for different values of k . As k increases the cooperation decrease for a given value of memory h and rewiring probability θ . In case of higher k , in a given round more agents update their type which means that agents have to learn the other's type more frequently and they will store more false information as memory gets large. The memory to reach positive cooperation seems to be unchanged for different values of k .

In the previous analysis the preference bias of CC type for cooperation has been set to $w_{CC} = 0.35$, which was the value by which the CC type agent cooperates if the opponent has been cooperating in half of the cases about which she has information.¹⁶ If we increase the value of w_{CC} , the CC type agents require less cooperative examples from the past to cooperate. As it can be seen on Figure 13, if $w_{CC} > 0.35$, the level of cooperation doesn't change much in case of $\rho = 2$, but if w_{CC} is too low, the cooperation stays at zero. This means that if the conditional cooperators need too much cooperative evidence to cooperate they end up defecting too many times. But they need to be permissive just up to a point for the emergence of cooperation, further indulgence doesn't improve the cooperation. In case of $\rho = 4$, however, cooperation increases with w_{CC} when it is higher than the middle value 0.35 (Figure 17).

Figure 14 and 18 show how the cooperation changes if we modify the value of λ , the weight which is used by the agents to evaluate the relative importance of own and of received information. It can be seen on the graph that the level of cooperation is maximized at $\lambda = 0.5$ for any given value of h and θ : when agents equally weight the two sources of information. The cooperation is the lowest when individuals give just a little importance to their own experience for both $\rho = 2$ and $\rho = 4$.

It can be seen on these graphs that the overall pattern of cooperation summarized by Result 1 and Result 2 is robust to these parameter changes in most of the cases if the parameters are in a range when cooperation can emerge at all.

To reinforce these casual observations we run logit regressions of cooperation on all the parameter values. The result are as expected based on the previous discussion, see Table 6 in the Appendix.

5 Conclusion

We analyzed the evolution of cooperation in a setting that comes very close to actual human interactions. Agents with limited memory interact and share information through a given social network. We have seen that cooperation can but need not evolve in such a setting. Our results show that if there is "Stigma" there exists an interior memory size maximizing the probability to observe cooperative outcomes in the long run. In the "No Stigma" case, on the other hand, a larger memory always leads to better chances for cooperation.

¹⁶This happens given the other two payoff parameters: $a = 0.8$, $d = 0.5$.

These results add an interesting new viewpoint to the discussion of why evolution did not endow us with unlimited brain capacity or, more loosely speaking, did not make us "infinitely" smart. Not only is it the case that larger brain or memory capacity involves a direct cost, our results also show that it can be optimal (at least from a population viewpoint) to be endowed with finite memory in strategic situations. In line with the group selection literature in Biology limited memory in our model can emerge from the conflict of two populations endowed with different memory capacity in the brain. Since human memory capacity is in reality limited, our results also suggest that stigma may play an important role in the evolution of human cognitive capacities.

Future research will answer the question of whether limited memory is also evolutionary optimal from an individual perspective. Preliminary results show that this seems to be the case.

References

- [1] Andreoni, J. and Miller, J.H. (1993): Rational cooperation in the finitely repeated Prisoner's Dilemma: Experimental evidence, *The Economic Journal* 103, 570-585.
- [2] Axelrod R. (1984): *The Evolution of Cooperation*, New York: Basic Books.
- [3] Axelrod R. and Hamilton, W.D. (1981). The evolution of cooperation, *Science* 211, 1390-1396.
- [4] Axelrod, R., Riolo, R.L., and Cohen, M.D. (2002): Beyond geography: Cooperation with persistent links in the absence of clustered neighborhoods, *Personality and Social Psychology Review* 6, 341-346.
- [5] Boyd, R. and Richerson, P.J. (1989): Evolution of Indirect Reciprocity, *Social Networks* 11, 213-236.
- [6] Cassar, A. (2007): Coordination and cooperation in local, random and small world networks: Experimental evidence, *Games and Economic Behavior* 58, 209-230.
- [7] Cortens, R. and K. Cook, *Cooperation and Reputation in Dynamic Networks*, working paper Utrecht University.
- [8] Cosmides, L. and Tooby, J. (1992): Cognitive adaptations for social exchange. In (eds.): Barkow, J.H., Cosmides, L., and Tooby, J.: *The Adapted mind: Evolutionary psychology and the generation of culture*, Oxford: Oxford University Press.
- [9] Cowan, N. (2001): The magical number 4 in short-term memory: A reconsideration of mental storage capacity, *Behavioral and Brain Sciences*, 24, 87-185.

- [10] Eshel, I., Binmore, K., and Shaked, A. (1998): Altruists, egoists, and hooligans in a local interaction model, *American Economic Review* 88, 157-179.
- [11] Fehr, E. and Gächter, S. (2002): Altruistic punishment in humans, *Nature* 415, 137-140.
- [12] Fischbacher, U. and Gächter, S. (2008): Heterogeneous social preferences and the dynamics of free riding. CeDEx Discussion Paper No. 2008-07.
- [13] Fischbacher, U., Gächter, S., and Fehr, E. (2001): Are people conditionally cooperative? Evidence from a public good experiment, *Economic Letters* 71, 397-404.
- [14] Fosco, C. and F. Mengel (2008), Cooperation through Imitation and Exclusion in Networks, working paper Maastricht University.
- [15] Fudenberg, D. and Maskin, E. (1986): The Folk theorem in repeated games with discounting or with incomplete information, *Econometrica* 50, 533-554.
- [16] Godfray, H.C.J. (1992): The evolution of forgiveness, *Nature* 355, 206-207.
- [17] Goyal, S. (2007): Connections: an introduction to the economics of networks. Princeton University Press.
- [18] Grimm, V. and Mengel F. (2008): Cooperation in Viscous Populations – Experimental Evidence, *Games and Economic Behavior*, forthcoming.
- [19] Kreps, D.M., Milgrom, P., Roberts, J., and Wilson, R. (1982): Rational cooperation in the finitely repeated Prisoner's Dilemma, *Journal of Economic Theory* 27, 245-252.
- [20] Kurzban, R. and Houser, D. (2005): Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations, *Proceedings of the National Academy of Sciences* 102, 1803-1807.
- [21] Leimar, O. and Hammerstein, P. (2001): Evolution of cooperation through indirect reciprocity, *Proceeding of the Royal Society B*, 268, 745-753.
- [22] Mashuda, N. and Aihara, K. (2003): Spatial prisoner's dilemma optimally played in small-world networks, *Physics Letters A*, 313, 55-61.
- [23] Mengel, F. (2007): The Evolution of Function-Valued Traits for Conditional Cooperation", *Journal of Theoretical Biology* 245, 564-575.
- [24] Mengel, F. (2008): Matching Structure and the Cultural Transmission of Social Norms", *Journal of Economic Behavior and Organization* 67, 608-623
- [25] Milinski, M., Semmann, D., Bakker, T.C.M., and Krambeck, H. (2001): Cooperation through indirect reciprocity: image scoring or standing strategy?, *Proceedings of the Royal Society B* 268, 2495-2501.

- [26] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- [27] Milinski, M. and Wedekind, C. (1998): Working memory constrains human cooperation in the Prisoner's Dilemma, *Proceedings of the National Academy of Sciences* 95: 13755-58.
- [28] Mohtashemi, M. and Mui, L. (2003): Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism, *Journal of Theoretical Biology* 223, 523-531.
- [29] Nakamaru, M., Matsuda, H., and Iwasa, Y. (1997): The evolution of cooperation in a lattice structured population, *Journal of Theoretical Biology* 184, 65-81.
- [30] Nakamaru, M., Nogami, H. , and Iwasa, Y. (1998): Score-dependent fertility model for the evolution of cooperation in a lattice, *Journal of Theoretical Biology* 194, 101-124.
- [31] Nowak, M.A., and Sigmund, K. (1992): Tit for tat in heterogeneous populations. *Nature* 355: 250-253.
- [32] Nowak, M.A., and Sigmund, K. (1998): The Evolution of Indirect Reciprocity, *Nature* 437, 1291-1298.
- [33] Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak M.A. (2006): A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502-505.
- [34] Ohtsuki, H. and Nowak M.A. (2007): Direct reciprocity on graphs. *Journal of Theoretical Biology* 247, 462-470.
- [35] Pollock, G. and Dugatkin, L.A. (1992). Reciprocity and the emergence of reputation, *Journal of Theoretical Biology* 159, 25-37.
- [36] Rubinstein, A. (1979): Strong perfect equilibrium in supergames, *International Journal of Game Theory* 9, 1-12.
- [37] Seinen, I, and Schram, A. (2004): Social status and group norms: Indirect reciprocity in a repeated helping experiment , *European Economic Review* 50, 581-602.
- [38] Stevens, J.R. and Hauser, M.D. (2004): Why be nice? Psychological constraints on the evolution of cooperation, *TRENDS in Cognitive Sciences* 8(2): 60-65.
- [39] Stevens, J.R. and Hauser, M.D. (2005): Cooperative brains: psychological constraints on the evolution of altruism. In: S. Dehaene, J.-R. Duhamel, M.D. Hauser, and G. Rizolatti (Eds.): *From Monkey Brain to Human Brain*. MIT Press, 159-187.

- [40] Ule, A. (2008): Partner Choice and Cooperation in Networks: Theory and Experimental Evidence, Springer.
- [41] Van Lange, P. A. M., Otten, W., De Bruin, E., and Joireman, J. A. (1997): Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence, *Journal of Personality and Social Psychology* 73, 733-746.
- [42] Vega Redondo, F. (2007): Complex Social Networks, Cambridge University Press.
- [43] Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393: 409–410.
- [44] Williams, G.C. (1986): Evolution Through Group Selection, Blackwell.
- [45] Wedekind, C. and Milinski, M. (2000): Cooperation through image scoring in humans, *Science* 288, 850-852.

6 Appendix

6.1 Proofs of Propositions 1 and 2

Proof of Proposition 1

Proof. We will show that there exists a number $K \in \mathbb{N}$ and a probability \hat{q} s.t. from any $s \in S$ the probability is at least \hat{q} to converge within K periods to an absorbing state. K and \hat{q} are time independent and state independent. Hence the probability of not reaching an absorbing state after at least nK periods is at most $(1 - p)^n$ which tends to zero as $n \rightarrow \infty$. Consider an arbitrary state $s(t)$. Wait for the next revision opportunity and assume wlg that it occurs at t . Denote by $i_{\max}(t)$ the player with the highest payoff $\pi_i(t)$ from her last interaction at t . Assume $\tau(i_{\max}(t)) = D$. ($\tau(i_{\max}(t)) = A; \tau(i_{\max}(t)) = CC$). Assume $i_{\max}(t)$ was last matched with a defector. Then all agents in the population must be either of type CC (choosing defection) or of type D . Assume not and denote with index k the players drawn to revise at t . There is positive probability that i_{\max} was not matched with any of the k players in her last interaction. Assume all players k are matched with i_{\max} for selection. Then, obviously at $t + 1 : \tau(k) = D, \forall k$. (i) Assume $i_{\max}(t)$ was last matched with an altruist. There is positive probability that all agents (including i_{\max}) will be matched with the same agents again during the next N periods. (ii) Assume $i_{\max}(t)$ was last matched with an conditional cooperator. Then assume that matching in the next N periods is unchanged but that instead of i_{\max} being matched two defectors are matched. This has positive probability. Iterate the matching described in (i) or (ii) until all agents except for possibly the agent last matched with i_{\max} have switched to defection (cooperation). Match this agent with i_{\max} for selection. Then all agents are choosing either cooperation or defection, implying that all agents will obtain the same per interaction payoff.

In finite time the reputation matrix will converge and we will have reached an absorbing state. Denote the probability with which this happens by $q_s > 0, \forall s \in S$ and let $\hat{q} = \min_{s \in S} q_s$. This completes the proof. ■

Proof of Proposition 2

Proof. Take any state $s(t)$ in which some agents cooperate and some defect. Since all agents have positive probability to be matched the event that a cooperator and a defector are matched and afterwards also paired for selection has positive probability. After such an event the cooperator will change her type, implying that $q(s, s) \neq 1$. ■

6.2 Example about α

Here we illustrate the effect of changing α on the matching probabilities for different networks. Let $N = 20$ and $\rho = 2$. When $\theta = 0$, each agent has 4 direct neighbors and 4-4 neighbors to 2, 3, 4 distance, plus 3 neighbors to 5 distance. The following table shows the matching probabilities for different values of α and different distance values.

$\alpha \backslash$ distance	1	2	3	4	5
0.5	0.109	0.066	0.04	0.024	0.015
1	0.16	0.059	0.022	0.008	0.003
1.5	0.194	0.043	0.01	0.002	0.0005
2	0.216	0.029	0.004	0.0005	0.0007
2.5	0.229	0.019	0.002	0.0001	0.0001

As α increases the matching with direct neighbors gets more and more likely event while the matching with distant agents becomes unlikely. When $\alpha = \infty$, agents are matched exclusively with their direct neighbors.

Rewiring the links introduces heterogeneity: agents will have different number of neighbors of different distances which generates heterogeneity in the matching probabilities. For example, if we apply rewiring with probability 0.1 on this network, we can pick two individuals labelled by A and B having the following "neighborhood structure":

A: 5 direct neighbors, 11 second order neighbors and 3 third order neighbors.

B: 2 direct neighbors, 4 second order neighbors, 6 third order neighbors, 7 fourth order neighbors.

Consequently, player A is going to be matched more frequently by other agents but she is going to play just once in each round (see the section about the interaction structure).

The following tables show how the matching probabilities of these two individuals change with α . For player A:

$\alpha \backslash \text{distance}$	1	2	3
0.5	0.078	0.047	0.029
1	0.106	0.039	0.014
1.5	0.132	0.029	0.007
2	0.153	0.021	0.003
2.5	0.169	0.014	0.001

For player B:

$\alpha \backslash \text{distance}$	1	2	3	4
0.5	0.122	0.074	0.045	0.028
1	0.216	0.079	0.029	0.011
1.5	0.306	0.068	0.015	0.003
2	0.375	0.051	0.007	0.0009
2.5	0.421	0.035	0.003	0.0002

It can be seen that the player with more heterogenous neighborhood structure (B) has higher probability to be matched with direct neighbors and her matching probabilities decay more with distance.

6.3 Regression tables

In the tables, the 5% significant coefficients are denoted with *. In parenthesis the t-statistics are shown. For $h > 10$ the memory variables are not significant individually, but they are jointly, as shown by the tests in the tables. Table 5 and 4 use the data generated using the basic parameter setting while Table 6 uses the data from the sensitivity analysis when the parameters had been changed.

Logit model, $\rho=4$	$\theta = 0$	$\theta = 0.01$	$\theta = 0.05$	$\theta = 1$	For all θ
dummy_h<=10	-6.18*	-7.78*	-13.86*	-	-10.29*
	(-9.35)	(-8.61)	(-5.17)	-	(-19.7)
h*dummy_h<=10	1.55*	1.94*	2.93*	-	1.64*
	(7.26)	(6.91)	(4.12)	-	(10.68)
h ² *dummy_h<=10	-0.11*	-0.13*	-0.17*	-	-0.11*
	(-6.61)	(-6.47)	(-3.74)	-	(-9.56)
dummy_h>10	1.6	3.51	0.98	-57.98*	-1.84
	(0.52)	(1.26)	(0.29)	(-2.78)	(-1.21)
h*dummy_h>10	-0.34	-0.57	-0.34	6.05*	-0.32
	(-0.8)	(-1.52)	(-0.75)	(2.54)	(-1.59)
h ² *dummy_h>10	0.006	0.014	0.008	-0.17*	0.008
	(0.46)	(1.17)	(0.55)	(-2.44)	(1.27)
Average Distance	-	-	-	-	-0.036
					(-1.15)
Clustering	-	-	-	-	5.26*
					(14.58)
Number of observations	2000	2000	2000	2000	11939
Log likelihood	-800.98	-794.96	-540.44	-120.38	-2875.60
Joint test of h>10 terms					
Test statistics	404.10	372.82	428.81	248.80	1131.99
P-value	0.00	0.00	0.00	0.00	0.00

Table 4: Logit regression for the case of $r=4$ and the basic parameter values.

Logit model, $\rho=2$	$\theta = 0$	$\theta = 0.01$	$\theta = 0.05$	$\theta = 1$	For all θ
dummy_h<=10	-6.86*	-8*	-10.68*	-26.51	-9.62*
	(-10.33)	(-10.63)	(-8.81)	(-1.14)	(-22.46)
h*dummy_h<=10	1.83*	2.2*	2.62*	4.4	1.93*
	(8.62)	(9.4)	(7.57)	(0.83)	(15.31)
h ² *dummy_h<=10	-0.11*	-0.14*	-0.16*	-0.2	-0.11*
	(-6.87)	(-8.09)	(-6.66)	(-0.67)	(-12.64)
dummy_h>10	1.66	-0.79	2.1	-4.78*	-1.7
	(0.78)	(-0.38)	(1.02)	(-2.01)	(-1.71)
h*dummy_h>10	-0.15	0.21	-0.25	0.37	0.04
	(-0.53)	(0.77)	(-0.91)	(1.21)	(0.31)
h ² *dummy_h>10	0.005	-0.009	0.008	-0.007	-0.0009
	(0.52)	(-0.97)	(0.89)	(0.45)	(-0.22)
Average Distance	-	-	-	-	0.022*
					(2.11)
Clustering	-	-	-	-	3.55*
					(13.09)
Number of observations	2000	2000	2000	2000	9262
Log likelihood	-1148.5	-1154.04	-1107.09	-652.13	-4986.51
Joint test of h>10 terms					
Test statistics	62.36	38.84	9.31	153.80	340.98
P-value	0.00	0.00	0.0255	0.00	0.00

Table 5: Logit regression for the case of $r=2$ and the basic parameter values.

Logit model	$\rho=2$	$\rho=4$
dummy_h<=10	-19.92* (-119.64)	-22.1* (-101.74)
h*dummy_h<=10	1.75* (46.23)	1.67* (35.92)
h ² *dummy_h<=10	-0.99* (-36.93)	-0.103* (-30.42)
dummy_h>10	-12.58* (-57.77)	-14.89* (-33.08)
h*dummy_h>10	0.039* (2.03)	-0.06 (-0.97)
h ² *dummy_h>10	-0.00005 (-1.75)	-0.0003 (-0.18)
α	1.15* (59.78)	1.35* (38.33)
w_{CC}	19.96* (96.46)	24.77* (99.86)
k	-0.03* (-58.68)	-0.07* (-39.35)
λ	2.48* (41.38)	0.83* (9.48)
Average distance	0.004 (1.25)	0.007 (0.71)
Clustering	4.35* (48.78)	5.88* (56.21)
Number of observations	157486	203245
Log likelihood	-62136.67	-33981.43

Table 6: Logit regression for all values of the parameters

6.4 Graphs of the sensitivity analysis

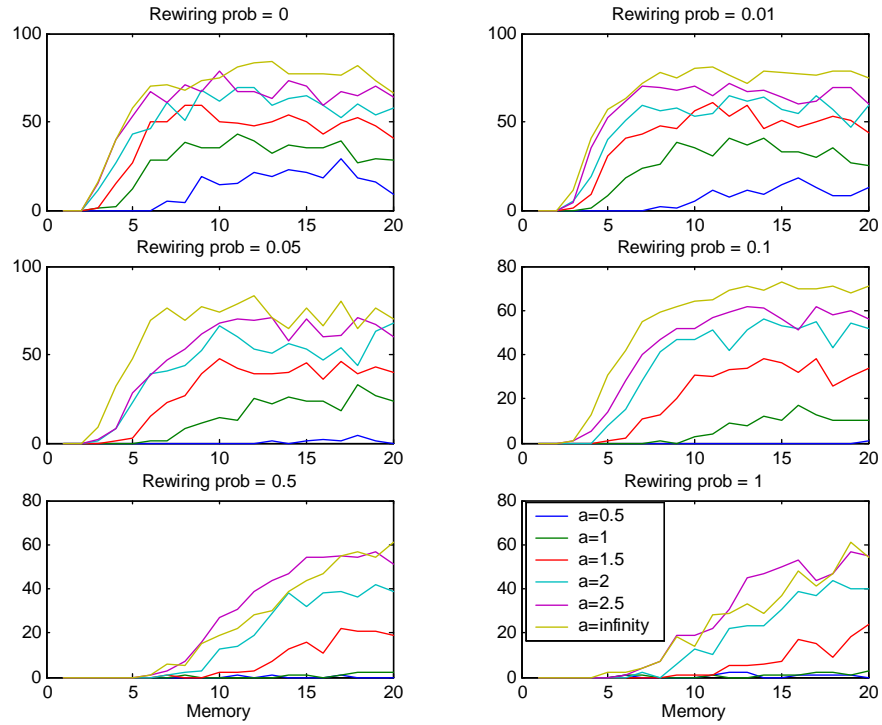


Figure 11: Frequency of cooperative outcomes as a function of memory for different values of θ and α when $\rho = 2$.

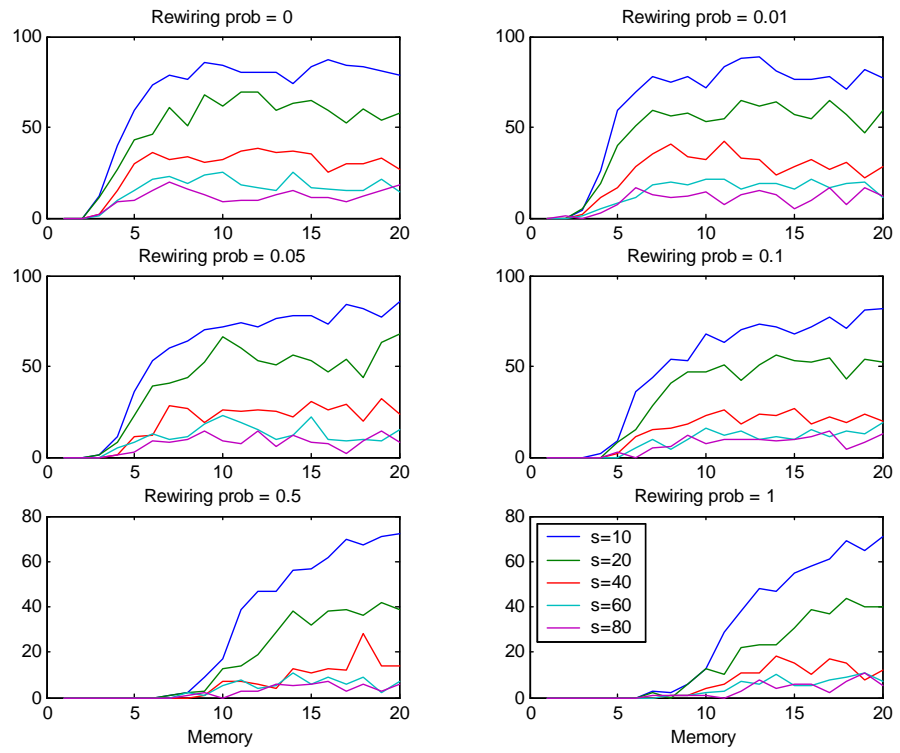


Figure 12: Frequency of cooperative outcomes as a function of memory for different values of θ and k for $\rho = 2$.

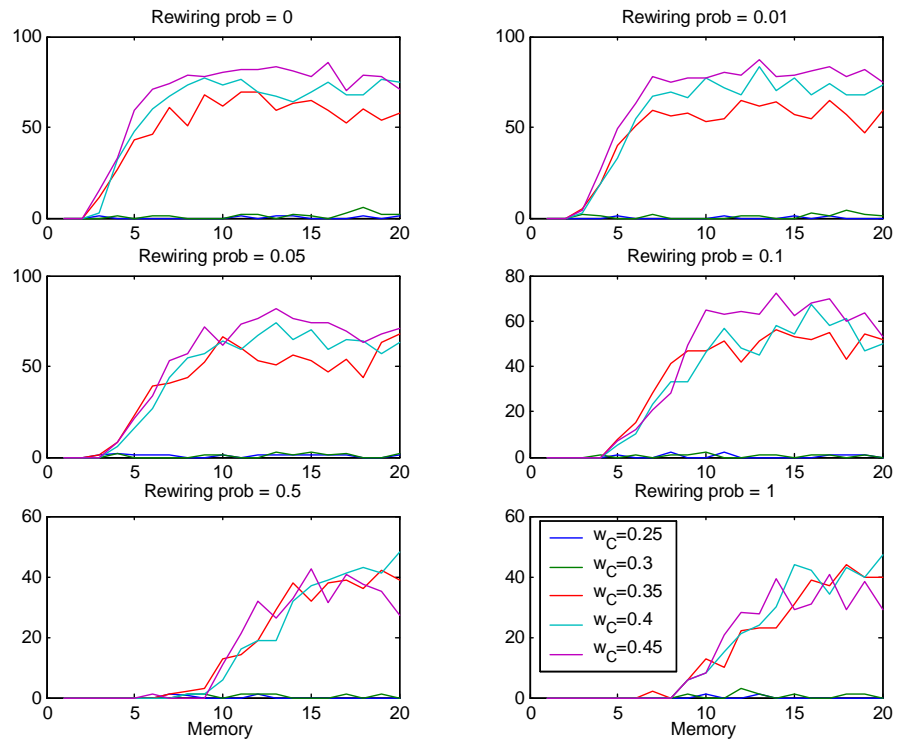


Figure 13: Frequency of cooperative outcomes as a function of memory for different values of θ and w_{CC} when $\rho = 2$.

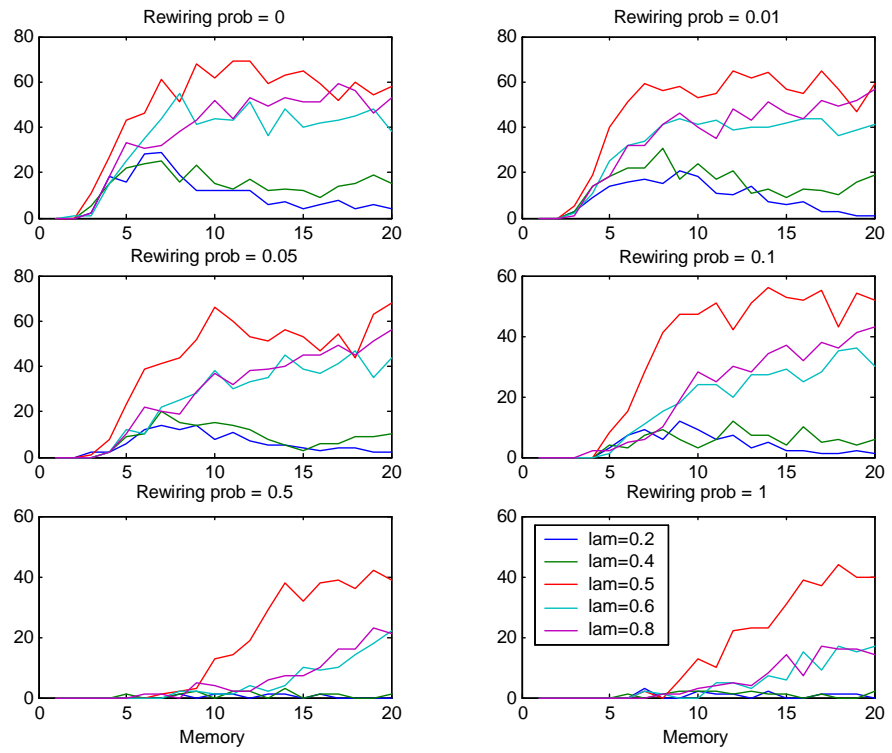


Figure 14: Frequency of cooperative outcomes as a function of memory for different values of θ and λ when $\rho = 2$.

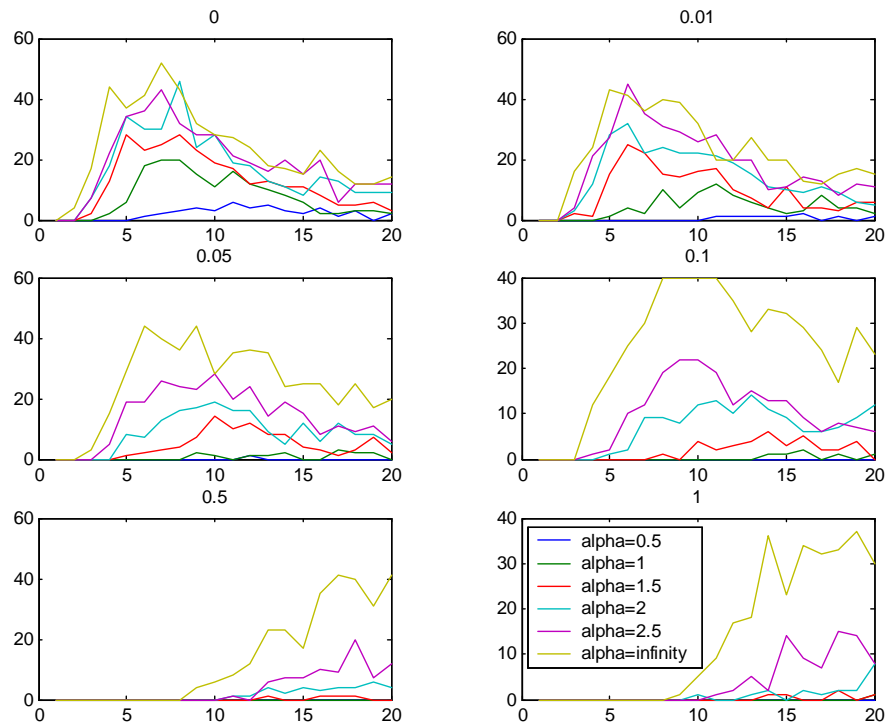


Figure 15: Frequency of cooperative outcomes as a function of memory for different values of θ and α when $\rho = 4$.

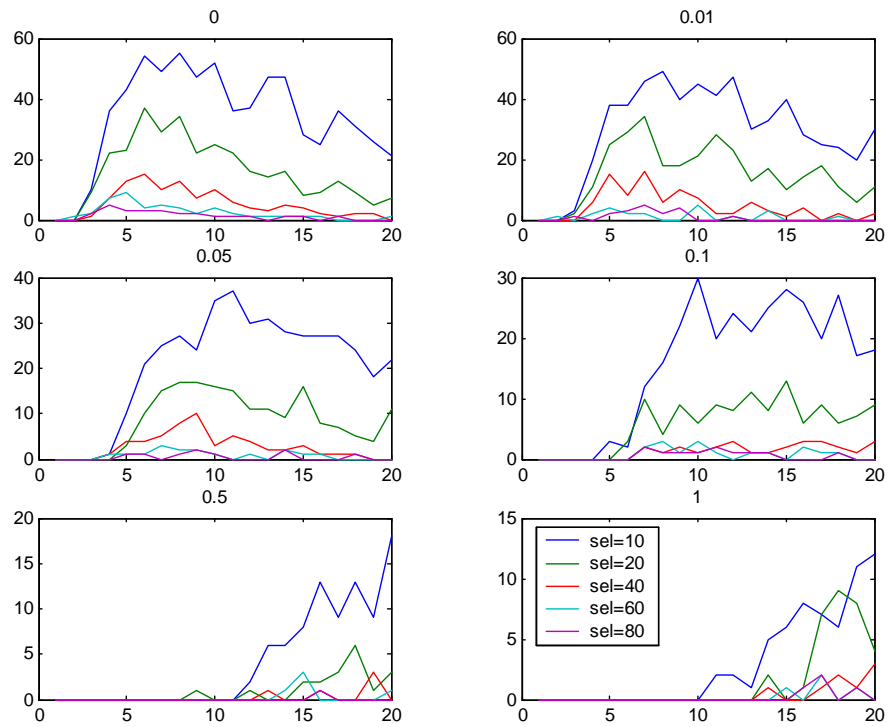


Figure 16: Frequency of cooperative outcomes as a function memory for different values of θ and k when $\rho = 4$.

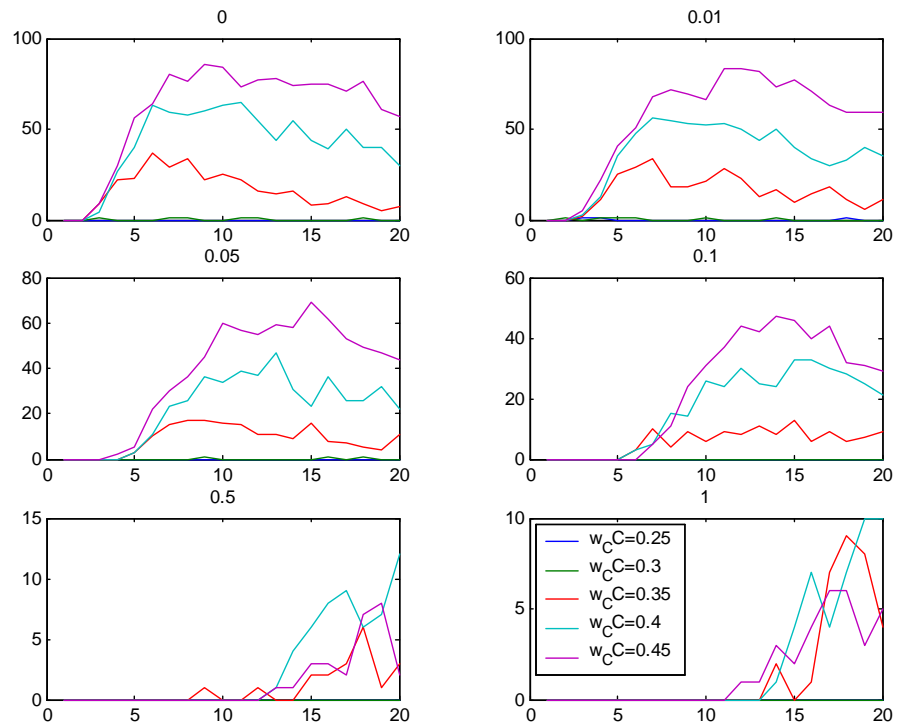


Figure 17: Frequency of cooperative outcomes as a function of memory for different values of θ and w_{CC} when $\rho = 4$.

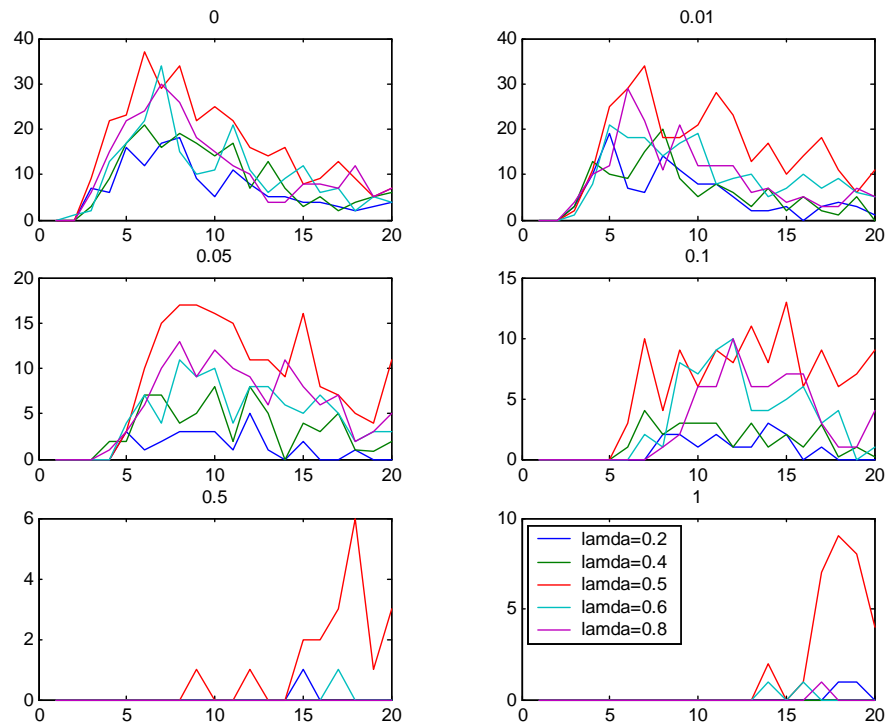


Figure 18: Frequency of cooperative outcomes as a function memory for different values of θ and λ when $\rho = 4$.