The option to leave: Conditional dissociation in the evolution of cooperation

Segismundo S. Izquierdo, Luis R. Izquierdo, Fernando Vega-Redondo

Author's Accepted Manuscript

The option to leave: Conditional dissociation in the evolution of cooperation

Segismundo S. Izquierdo, Luis R. Izquierdo, Fernando Vega-Redondo

www.elsevier.com/locate/yjtbi

Cite this article as: Segismundo S. Izquierdo, Luis R. Izquierdo and Fernando Vega-Redondo, The option to leave: Conditional dissociation in the evolution of cooperation, *Journal of Theoretical Biology*, doi:10.1016/j.jtbi.2010.07.039

The Option to Leave: Conditional Dissociation in the Evolution of Cooperation

Segismundo S. Izquierdo [a], Luis R. Izquierdo [b], Fernando Vega-Redondo[c]

[a]Department of Industrial Organization, University of Valladolid, 47011 Spain; InSiSoc. [b]Department of Civil Engineering, University of Burgos, 09001 Spain; InSiSoc. [c]European University Institute, I-50133 Florence, Italy; Instituto Valenciano de Investigaciones Económicas, 46020 Valencia, Spain

Corresponding Author:

Segismundo S. Izquierdo

Departamento de Organización Industrial,

ETS Ingenieros Industriales,

Pº del cauce s/n

47011 Valladolid, Spain

Phone: +34 983 18 38 73

Fax: +34 983 42 33 10

e-mail: segis@eis.uva.es

Luis R. Izquierdo   e-mail: luis@izquierdo.name

Fernando Vega-Redondo   e-mail: Fernando.Vega@eui.eu

Abstract

Conditional dissociation, i.e. the option to leave an interacting partner in response to his behaviour, is a mechanism that has been shown to promote cooperation in several settings, but the fundamental features that make conditional dissociation work in this way are not yet fully understood. This paper identifies some of the key conditions that make conditional dissociation lead to high levels of cooperation, explains how this mechanism can support the evolutionary coexistence of cooperative and non-cooperative behaviour typically observed in nature, and provides an analytical formula to estimate the expected degree of cooperation thus achieved. Our model involves a population of individuals who are paired to play an iterated prisoner's dilemma. All individuals share the same capacity to react to the action previously chosen by their partner and, without any other *a priori* constraint or exclusion, they may use *any* behavioural rule that is compatible with this capacity. The dynamic evolution of the population eventually enters either a non-cooperative or a partially cooperative regime, depending mainly on the expected lifetime of individuals. Whenever the partially cooperative regime materializes, the cornerstone of its long-run stability is the coexistence of defectors and "Out-for-Tat" strategists, the latter being those who start cooperating and respond to defection by merely leaving. We find, therefore, that conditional dissociation is the essential disciplinary device supporting cooperation, whilst other conditional strategies (such as Tit-for-Tat) remain present only in small population shares. These conclusions are obtained both by extensive numerical simulations and through analytical mean-field methods that approximate the stochastic simulation dynamics and deliver accurate predictions for general parameter configurations.

2

Keywords: Evolutionary dynamics, Prisoner's Dilemma, Positive assortment,

Conditional association, Exit option.

Introduction

The evolutionary emergence of cooperation among unrelated individuals is still a fundamental puzzle that keeps fascinating evolutionary biologists and social scientists (Mathew and Boyd, 2009; McNamara et al., 2008; Nowak, 2006; Sachs, 2006). Indeed, "How did cooperative behaviour evolve?" is considered to be one of the top 25 big questions faced by modern science (Pennisi, 2005). A unifying framework that has been put forward to explain how altruism may emerge in an evolutionary context is the existence of *positive assortment* mechanisms: processes or conditions that allow cooperators to preferentially interact among themselves (Doebeli and Hauert, 2005; Eshel and Cavalli-Sforza, 1982; Fletcher and Doebeli, 2006; Fletcher and Doebeli, 2009; Fletcher and Zwick, 2004; Le and Boyd, 2007; Németh and Takács, 2010; Nowak et al., 2010; Taylor and Day, 2004).

Pre-interaction partner-selection or partner-refusal is one of the possible mechanisms for positive assortment. However, this conditional association mechanism requires that individuals have the capacity to recognize other individuals and anticipate their behaviour. Past experience (Aktipis, 2006; Ashlock et al., 1996; de Vos et al., 2001; Janssen, 2008; Vanberg and Congleton, 1992), external signs and tags (Edmonds and Hales, 2003; Hales and Edmonds, 2003; Jansen and van Baalen, 2006; McDonald and Sen, 2006; Riolo et al., 2001; Tanimoto, 2007) or a projection of the agent's own intentions onto his potential partner (Macy and Skvoretz, 1998; Orbell and Dawes, 1991) are alternative ways proposed in the literature to explain how such anticipation may take place.

A simpler mechanism that can also promote positive assortment among individuals in repeated interactions is the option to part ways and look for a new partner if unhappy with the current one, i.e. *post-interaction* partner-refusal (Aktipis, 2004; Axelrod, 2000; Schuessler, 1989; Thomas and Feldman, 1988). In computer simulations, this mechanism has been shown to promote cooperation even when new associations are random (Schuessler, 1989). To emphasize the difference between *pre-interaction* selection of a *new* partner and *post-interaction* refusal of the *current* partner, we will refer to the latter as *conditional dissociation*. Conditional dissociation does not require complex cognitive capabilities. It requires neither memory nor the ability to anticipate the behaviour of new partners. It only presumes the capacity to escape an unpleasant situation, which is indeed a trait that has evolved in most organisms in nature (Rowley, 1983; Simms et al., 2006). Furthermore, in contrast with, say, recognition and assortment based on signals and tags, conditional dissociation is based on observed behaviour and thus has the advantage of being cheating-proof. And finally, the option to leave seems to be very common in nature; in the real world, as frequently pointed out (Macy and Skvoretz, 1998; Orbell and Dawes, 1993; St-Pierre et al., 2009), many "prisoner's dilemmas" rarely involve "prisoners": instead, partners are usually free to walk away.

A number of studies, both computational and experimental (Coricelli et al., 2004; Hauk and Nagel, 2001), suggest that conditional dissociation is a factor that can enhance cooperation in social dilemmas. Some studies assume immediate re-coupling after leaving a partner (Schuessler, 1989) whilst others allow for an intermediate state of non-interaction in which individuals receive payoffs from some alternative activity (Aktipis, 2004; Hauk, 2003; Vanberg and Congleton, 1992). The exit option has also

been studied in contexts where the pattern of interaction takes the form of a social network (Pacheco et al., 2006; Santos et al., 2006; Szabó and Hauert, 2002), in public good games (Hauert et al., 2002a; Hauert et al., 2002b; Hauert et al., 2007; Mathew and Boyd, 2009) and in continuous-action games in which the threshold for maintaining an ongoing partnership coevolves with the level of cooperation (McNamara et al., 2008; Sherratt and Roberts, 1998).

A common approach to study the effect of conditional dissociation has been to confront the strategy that cooperates and quits after a partner's first defection – a strategy that has been called CONCO (Schuessler, 1989; The EdK-Group, 2000), Out-for-Tat (Hayashi, 1993), MOTH (Joyce et al., 2006) and Walk Away (Aktipis, 2004) – with other strategies that either seem well suited to the context in hand (Yamagishi et al., 1994) or have proven successful in other environments, such as Tit-for-Tat (TFT) (Axelrod, 1984; Axelrod, 2000) or Pavlov (Nowak and Sigmund, 1993). Hayashi and associates (Hayashi, 1993; Jin et al., 1996; Yamagishi et al., 1994) showed that conditional-dissociation strategies perform well even when a limited form of partner pre-selection is allowed, and Aktipis (2004) showed the robustness of Walk Away versus TFT and Pavlov in a spatial framework.

A primary concern with this line of research, however, is that conditional dissociation is typically combined with other features that are well known to promote the emergence of cooperation on their own, such as spatial or network effects (Aktipis, 2004), partner pre-selection (Hayashi, 1993; Hruschka and Henrich, 2006; Jin et al., 1996; Vanberg and Congleton, 1992; Yamagishi et al., 1994) or segregation in groups (Hamilton and Taborsky, 2005). This combination of factors makes it difficult to distil

the effective role played by conditional dissociation on the evolution of cooperation. An additional concern derives from the fact that the particular choice of competing strategies has a crucial bearing on the results (Bendor and Swistak, 1995), which calls for some underlying framework that justifies the particular ecology of strategies under consideration. Finally, the study of conditional dissociation has been carried out mainly through computer simulations and experimental studies; in general, it had not been possible to provide more analytical support for the observed results yet. A notable exception is provided by Fujiwara-Greve and Okuno-Fujiwara (2009), who conduct a static analysis of a model with the same sequence of events as the one presented in this paper, but with a different strategy space: the infinite set of all possible pure strategies that depend on the potentially infinite history of actions that may occur within each match.

To address these concerns, the present paper formulates a very basic setup – in a sense, a "minimalist" one – that nevertheless captures the most essential aspects of the question in hand, i.e. the significance of conditional dissociation in the evolution of cooperation. In our model individuals are randomly matched in pairs and remain together until one of them dies or decides to leave; their strategies govern both whether they cooperate or defect and whether they stay or not with their current partner – these are the two sole dimensions of choice. Naturally, they can make these choices dependent on past experience, and again we make a simple assumption in this respect: only the partner's most recent behaviour matters. Importantly, within such a stylized decision framework, we impose no restrictions and allow for all possible strategies. A natural way to motivate this approach is through the assumption that all individuals

enjoy the same (e.g. genetically-based) capacity to process and react to the information

contained in their own past experience (Imhof and Nowak, 2010).

By now it is well understood that if cooperation is to emerge in a large-population

context, some extent of endogenous positive assortment (in the general sense of the

term) must take place (Fletcher and Doebeli, 2009). Within our simple basic framework,

the only way in which this assortment may come about is through the effect that

conditional behaviour has in shaping partner's behaviour. Thus, in this respect, our

model provides only two possible routes for positive assortment: conditional

dissociation or conditional play. Which of the two represents the essential mechanism?

Our analysis provides a clear-cut answer: conditional dissociation – implemented by

cooperative Out-for-Tat strategies – is the dominant force at work, outcompeting

alternative strategies that rely on conditional play (such as Tit-for-Tat).

In our evolutionary process the population is in continuous flux and even in the

long run all possible strategies are represented. Nonetheless, we find clear regularities in

this dynamic stochastic process. As a matter of fact, we identify two, and only two,

dynamic regularities that together suffice to characterize the overall long-run dynamics

of the whole population. We call these dynamic patterns the cooperative and the

uncooperative regimes, and each of them is associated to a well-defined level of

aggregate cooperation (which is essentially zero in the uncooperative regime). Our

focus in this paper is to understand the emergence of the cooperative regime, and to

fully characterize how its associated level of cooperation depends on the key parameters

of the model. In this endeavour, we first resort to extensive numerical simulations,

which already provide a very sharp account of the situation. This is then complemented

8

by a mean-field analysis of the model, leading to a closed-form approximation of the level of cooperation in the general case. In fact, these analytical results are proven to be in very accurate correspondence with our simulation results, and extend our main conclusions to general parameter configurations.

The rest of the paper is organized as follows. First we describe the model. Next, we present and discuss the simulation results. And finally, we provide the mean-field analysis of the model and the formulas for the level of cooperation. Much of the technical details on both the simulations and the theoretical analysis are contained in the supplementary information.

The model

The population. We consider a population of N individuals who are matched in pairs to play a Prisoner's Dilemma game. To focus on the key issue at hand, we assume in the main body of the paper that N is constant and even, but this assumption can be relaxed. In particular, we show in the supplementary material that our qualitative and analytical results remain equally valid if the process is such that in every period each individual breeds one offspring with probability β, and may die with the same probability β (both events being stochastically independent), which induces a null expected growth but a variable population size.

The game. The two possible actions in the Prisoner's Dilemma game are **C** (for Cooperate) and **D** (for Defect). If both individuals cooperate, each obtains R (Reward); if both defect, each obtains P (Punishment); if one cooperates and the other defects, the cooperator gets S (Sucker) and the defector gets T (Temptation). The payoffs satisfy the

condition $T > R > P > S \geq 0$. Therefore, mutual cooperation is better than mutual defection for both individuals, but **D** is a dominant strategy in the one-shot game.

After each interaction, both individuals simultaneously decide whether to break their current partnership (Leave) or not (Stay). The Leave/Stay decision may be conditioned on the current partner's preceding action (**C** or **D**), hence the name "conditional dissociation". If any of the two individuals in a pair decides to leave, the partnership is broken and both individuals become single.

The strategies. Each individual carries three "switches" or "genes" that shape his behaviour. The first gene determines the individual's action, **C** or **D**, in the first interaction of any new partnership. The second gene determines what to do after the interaction if one's partner has played **C**. The options are labelled **C**, **D**, and **L**, where **L** stands for the decision to leave one's current partner, **D** stands for the decision to remain with one's partner and defect in the next period (assuming the partner does not leave or die), and **C** stands for the decision to remain with one's partner and cooperate in the next period (again, assuming the partner does not leave or die). Finally, the third gene determines the choice – also **C**, **D** or **L** – after one's current partner has just played **D**. A strategy thus specifies a course of action at every possible juncture in an individual's life.

We use the notation $X_1\_X_2\_X_3$ for strategies, where $X_1 \in \{\mathbf{C}, \mathbf{D}\}$ is the action to take with a new partner, $X_2 \in \{\mathbf{C}, \mathbf{D}, \mathbf{L}\}$ is the action to take if one's current partner cooperated, and $X_3 \in \{\mathbf{C}, \mathbf{D}, \mathbf{L}\}$ is the action to take if one's current partner defected. We permit all 18 possible strategies or decision rules. Hence, there are 171 strategy pairs.

The time line. The population process advances in discrete time periods $t = 0, 1, 2,....$ In each period, the following events occur in sequence (Fig. 1):

[Figure 1]

(i) *Random matching of singles*. At the beginning of each period, single individuals are randomly matched in pairs (partnerships), with all pairings being equally likely. At the beginning of period $t = 0$, all individuals are single.

(ii) *Pairwise interaction*. All pairs play the Prisoner's Dilemma game once. Each individual's action (**C** or **D**) is determined by his decision rule, as applied to his immediate past experience, i.e., whether she is involved in a new partnership or not, and, in the latter case, the action that his partner took in the preceding period.

(iii) *Conditional dissociation.* Every individual decides whether to leave or stay, according to his decision rule. Individuals involved in partnerships that are broken join the pool of singles entering the next period.

(iv) *Death*. Individuals randomly die, an event that occurs independently for each individual with probability β in each time period. Hence, the lifespan of an individual is geometrically distributed with mean value $\alpha = 1/\beta$. The death of an individual renders his partner single, so those individuals whose partner died also join the pool of singles entering the next period.

(v) *Replacement*. Each individual who dies is replaced by a fresh entrant, thus the population size remains constant across all time periods. The decision rule of each new entrant is assigned with probability proportional to each rule's aggregate payoff in the current period. The adoption of decision rules by new entrants is subject to independent random mutation, which occurs with constant probability μ. In this event, the entrant picks up a decision rule at random, all 18 decision rules chosen with equal probability.

Naturally, the entrants who replace dying individuals at any period are added to the starting pool of singles of the next period. Thus, this pool of singles is composed of new entrants, separated individuals from broken pairs, and "widowed" individuals (those whose partner died). Note that, for N even, the number of singles is always even.

It is worth remarking that what has been called "death" and "replacement / birth" can be interpreted literally as biological reproduction proportional to payoffs or, alternatively, in terms of learning by imitation. In the latter interpretation, entrants copy the decision rules of individuals in the current population independently, in a way such that the particular decision rule of any given individual is copied with a probability proportional to that individual's payoff in the current period. The parameter β should then be conceived as the probability of a "learning draw" and μ as the probability of mistakes and/or experimentation (Binmore et al., 1995; Björnerstedt and Weibull, 1996).

The above sequence of events defines a finite Markov chain over discrete time $t = 0, 1, 2,...$ For β, $\mu > 0$, any state can be reached from any other state with positive probability and the transitions are aperiodic. This implies that the process is ergodic, i.e. there is a unique invariant distribution to which it converges from any initial distribution. An important consequence of all this is that the long-run fraction of the time that the system spends in each state is strictly positive and independent of the initial conditions.

Simulation results

For the sake of focus, we present here simulations with payoff values $T = 4$, $R = 3$, $P = 1$, $S = 0$. The benchmark case will be a population of $N = 1000$ individuals with a

12

mutation rate of $\mu = 0.05$. All the numerical experiments in this paper can be replicated using the available computational model (see supplementary material). In terms of the observed outcomes of the process in any given period, let x_{DD} be the share of DD outcomes and let x_{CC} be the share of CC outcomes, namely, the *level of cooperation*. For any given value of the individuals' expected lifespan, α, the system evolves towards one of the following two regimes:

i) A partially cooperative regime in which the time series of the level of cooperation x_{CC} remains within a narrow band around a significantly positive particular level $x_{CC}^{CR} > 0.3$. Specifically, we denote by $\hat{x}_{CC}^{CR}$ the estimated average level of cooperation in that narrow band (see Table S1 for details). In the reported simulations we will consider the system to be in the cooperative regime if $| x_{CC} - \hat{x}_{CC}^{CR} | \leq 0.1$.

ii) A non-cooperative regime characterized by levels of DD outcomes x_{DD} close to 1. In the reported simulations we will consider the system to be in the non-cooperative regime if $x_{DD} \geq 0.8$.

The results crucially depend on individuals' expected lifespan, $\alpha = 1/\beta$, i.e. the expected number of interaction periods in an individual's life. Thus, in accord with intuition (see the continuous Prisoner's Dilemma model of McNamara et al. (2008) for a similar phenomenon), we find that as α grows, the system switches from a regime dominated by defection within pairs (DD outcomes) to a regime characterized by a high level of cooperation (CC outcomes). This level of cooperation also increases as α grows. To understand this behaviour in detail, it is useful to organize the discussion into three broad scenarios: short, long, and intermediate expected lifespan α.

13

(I) *Short life*: For low values of α, the system quickly evolves towards the non-cooperative regime. In particular, for $\alpha = 5$, Fig. 2 shows the evolution of x_{DD} in 18 different runs, each run starting with the whole population using one of the 18 possible pure strategies. In all these cases, the system reaches the non-cooperative regime in less than 1000 stages, even when the initial population is made up entirely by cooperative strategists.

[figure 2]

The behaviour displayed in Fig. 2 is typical for low values of α (Table S2). In those cases, the unique regime observed after a short initial transient phase is always the non-cooperative regime. Specifically, for $\alpha = 5$ we find that, after some periods of adjustment, the average fraction of DD outcomes, x_{DD}, is above 80% in more than 99.99 % of the periods (Table S2). Besides, the main strategies observed in this non-cooperative regime (commanding more than 90% of the population frequencies in total) are those that start any new partnership with defection and respond to defection by either defecting or leaving (Table S3).

(II) *Long life*: If the expected life of individuals is long enough, the behaviour of the process is drastically different: the system evolves towards a cooperative regime characterized by high levels of mutual cooperation, x_{CC}. In particular, for $\alpha = 100$, Fig. 3 shows the evolution of x_{CC} in 18 different runs (again each run starts with the whole population using one of the 18 different strategies). In every case, in less than 4000 stages the system reaches a cooperative regime in which x_{CC} remains around $\hat{x}_{CC}^{CR} = 82.7\%$. This behaviour is indeed typical for such a high value of α (Table S2).

14

[Figure 3]

Concerning the strategy composition (Table S4), the results show a clear predominance of the **C_C_L** strategy (Out-for-Tat; 77.5%), distantly followed by the **D_X_L** family (where **X** stands for **C**, **D** or **L**; 6.9%) and by **C_C_D** (Tit-for-Tat; 6.5%). Given this distribution of strategies, we note that protection against defectors is mainly achieved by leaving such partners, while cooperating pairs tend to remain together. This clearly illustrates the evolutionary emergence of positive assortment, i.e. evolutionary pressures act in a way such that cooperators tend to interact among them more than expected based on population averages.

(III) *Intermediate life*: For intermediate values of α, the situation is more complex than in the former two scenarios. Fig. 4 illustrates the situation for an expected lifespan $\alpha = 20$. We observe that, for most initial conditions, the system is first attracted towards the non-cooperative regime (thus, in a sense, this regime can be conceived as enjoying a larger "basin of attraction"), and stays there for some time. But thereafter a transition towards the more persistent cooperative regime eventually occurs.

[Figure 4]

Given our window of observation, to encounter the transitions across both regimes implied by ergodicity we must consider other parameterizations (e.g. smaller populations) where, indeed, frequent jumps between those two regimes are observed (see supplementary figures). In general, of course, which regime is reached first depends probabilistically on the initial state. However, as the expected lifespan increases, so does the "basin of attraction" of the cooperative regime, as well as its persistence and its estimated reference level of cooperation $\hat{x}_{CC}^{CR}$ (Fig. 5, Table S2).

15

Importantly, the results on Fig. 5 (as well as Fig. 3, Fig. 4 and Table S2) starkly show that, for all values of α, the dynamics of the process gravitate in one of the two regimes that we have defined. (To be precise, the system spends at least 97.5 % of the time in the limited subset of states defined by the union of the two regimes).

[Figure 5]

Next, in order to see how the value of α impinges on strategy composition, refer to Fig. 6, which provides the composition of strategies when the system is in the cooperative regime (see Table S4 for the specific values and standard deviations). A first remarkable observation is that the same strategies dominate the cooperative regime in all cases, both when the value of α is as low as 11 and when it is as high as 100. In particular, the Out-for-Tat strategy **C_C_L** that reacts to defection by leaving the partner in question is by far the one with the highest frequency. This predominance is reinforced as α grows, with the main contender for high values of α being Tit-for-Tat (whose average frequency nonetheless remains below 7%).

[Figure 6]

The dominant prevalence of Out-for-Tat over TFT is particularly interesting given the well-known fact that TFT can sustain cooperation in environments where leaving is not an option. Indeed, if we remove the option to leave from our particular framework by considering only the 8 strategies that always stay with their partner, and we assume a sufficiently long expected life, then high levels of cooperation supported by TFT are naturally achieved (see Fig. 7). As an example, using the same simulation conditions as in Fig. 5 and 6, for expected life $\alpha = 25$, we obtain an average level of mutual cooperation $x_{CC} = 62.7\%$ (s.d. 2.2%) and an average frequency of TFT of 58.6%

16

(s.d. 2.1%). These results contrast with those obtained when the option to leave is allowed and TFT has to compete with Out-for-Tat: in this latter case, we obtain an average level of mutual cooperation x_{CC} = 70.2% (s.d. 3%), and the observed average frequency of TFT is 4.7%, versus an average frequency of 69.8% for Out-for-Tat.

[figure 7]

Discussion of simulation results

How can we understand the coexistence of such diverse strategies in the partially cooperative regime? To answer this question we have to fully grasp the mechanism by which positive assortment emerges and prevails in this cooperative regime, i.e. the reason for which the strategy composition in the set of couples is entirely different from the strategy composition in the pool of singles, and this systematic difference persists over time. Simplifying to some extent, the set of stable couples is mostly composed of happily paired Out-for-Tat strategists, whilst the pool of singles is predominantly made up of avid defectors that live off the new Out-for-Tat singles (widowed, separated or newborn) by exploiting them in unstable one-period partnerships until two Out-for-Tats meet each other. Once the cooperative partner is found, the stable couple of Out-for-Tats will live happily together until death do them part. Thus, the observation that the level of cooperation increases with the expected lifetime, α, is not surprising. The longer the lifetime, the greater the chances Out-for-Tats have to find each other, and the longer time they have to reap the benefits of cooperation. Conversely, in the extreme case of α = 1 (one-period lifetime), there is no room for endogenous assortment and defectors

always do better than cooperators, so the evolutionary emergence of positive assortment cannot occur.

So, why do Out-for-Tat strategists not overtake the whole population? Why is the cooperative regime only partially so? The reason is that a population of Out-for-Tats is initially invadable by defectors. In particular, to play **D** and then leave gives a reasonably high payoff *if* many of the other singles initially play **C**. Indeed if all these other singles played **C**, then a single individual using the "hit and run" strategy **D_L_L** would earn the maximum payoff *T* in all periods and his genes would rapidly spread over the population. This explains why the cooperative regime is only partially so: in the cooperative equilibrium Out-for-Tats coexist with some **D**-players. And the reason why these defectors do not eventually dominate the population is that most of them are repeatedly sent to the pool of singles, so they interact quite frequently with each other and therefore earn only moderate payoffs. On the other hand, cooperators, through conditional dissociation, are able to develop a mechanism of positive assortment that allows them to avoid frequent exploitation and thus earn moderately high payoffs as well. In the end, the expected payoff to all strategies with a significant presence is approximately the same, as required for a steady state.

These clear-cut simulation results are particularly interesting, especially given our *unbiased* set of possible strategies. Note that, in this model, ergodicity implies that the system spends a strictly positive fraction of time in all possible states, i.e. sooner or later every state of the system is visited in any given run. However, our results show that the system spends practically all the time in a very narrow subset of states –those corresponding to the two regimes– and, as α grows, only the cooperative regime is observed in practical terms. This regime is polymorphic and dynamic: individuals with

18

all possible strategies come and go, but evolutionary forces make positive assortment emerge and be preserved, even though it is being constantly put to the test by mutants that may adopt any possible strategy within our unbiased set of behavioural complexity.

A mean-field approximation

Having understood the mechanism by which conditional dissociation can lead to a partially cooperative regime, the next step is to conduct a full-fledge mathematical analysis to confirm our conclusions for general values of the payoffs. In particular, this section presents a simple analytical formula that accurately approximates the level of cooperation in the cooperative regime. Recall that the stochastic process induced by our model can be represented as a time-homogeneous Markov chain with finite state space. The problem with this approach is that the state vector for this process needs to keep track of the assortment of strategies at any given period (171 variables) and of some additional information from the previous period. Thus, the number of possible states for a population of N individuals is greater than $\binom{170+N/2}{170}$. As an example, assuming $N = 400$, the number of possible states greatly exceeds the estimated number of atoms in the universe. Clearly this approach is not feasible.

In our journey towards simplicity, a next step would be to average out stochasticity by constructing the mean-field approximation of the stochastic process. This is a time-homogeneous system of ordinary differential equations (ODEs), one for each state variable, describing its expected law of motion (Benaim and Weibull, 2003). In our case this system involves more than 200 equations and is still intractable in practical terms. If we are interested in characterising the cooperative equilibrium associated with the cooperative regime, a possible approach that keeps the essential

19

forces at work in that regime and at the same time is manageable in analytical terms, consists in building a reduced system of ODEs considering only a subset of strategies. Naturally, the selected strategies are those that our numerical simulations have singled out as the most prevalent in the cooperative regime. Specifically, we work with a reduced system that includes the cooperative Out-for-Tat strategy **C_C_L** and all the strategies of the form **D_X_L** (where **X** is allowed to take all three possible values: **C**, **D** or **L**). For the outcomes of this reduced system, one may formulate a system of ODEs in just three state variables (x_{CC}, x_{DD} and x_{CD}) that represent the shares of interactions in which, respectively, both individuals play C, both play D, or each plays a different action (see supplementary information for a detailed derivation of the mean field and all the results in this section).

Fig. 8 depicts the unit simplex for the state vector $x = (x_{CC}, x_{DD}, x_{CD})$ of the simplified mean-field dynamics, and shows some orbits for the case considered in the simulations ($T = 4$, $R = 3$, $P = 1$, $S = 0$), under $\mu = 0$ and different values of α. For expected life $\alpha \leq 7$, the system has only one asymptotically stable critical point at $x_{DD} = 1$. By contrast, for $\alpha \geq 8$ there is also an interior asymptotically stable critical point that represents a partially cooperative equilibrium. We show below that this equilibrium neatly corresponds to the reference point defining the cooperative regime in our earlier discussion. Naturally, we observe that, as α increases, so does x_{CC} in the stable cooperative equilibrium.

[Figure 8]

Solving for the interior critical points in this case, we find that there is a lower bound for α, $\alpha_{Min} = 7.46$, that defines the region for which the interior critical points

appear. Provided the expected life α is above this critical value, and writing δ for $\delta = (1 - (1/\alpha))^2$, the corresponding level of cooperation at the critical points is given by

$$x_{CC}^* = 1 - \frac{1}{2\delta} \pm \sqrt{1 - \frac{3}{4\delta}}$$

where the solution in which the root has a positive sign corresponds to the (unique) *stable* cooperative equilibrium. This level of cooperation at the stable cooperative point x_{CC}^* can also be computed for an arbitrary mutation rate μ. Fig. 9 compares the analytical approximation x_{CC}^* for the reduced system with the average fraction of cooperating pairs $\hat{x}_{CC}^{CR}$ that characterizes the cooperative regime in our numerical simulations for the complete non-reduced system. The comparison is reassuring since, for different values of the individuals' expected life and mutation rates, we find that the theoretical predictions of the reduced system provide a very good approximation for the empirical values obtained in the non-reduced system.

[Figure 9]

Let us now turn to the general case of a prisoner's dilemma with arbitrary payoffs and, for simplicity, no mutation. Without loss of generality, we normalize the payoffs $[S, P, R, T]$ to $[0, p, r, 1]$, where $0 < p = \frac{P - S}{T - S} < r = \frac{R - S}{T - S} < 1$. With this normalization, and considering the reduced system to calculate the stable cooperative equilibrium, one can find both the minimum value of the expected life (α_{Min}) that is required for this cooperative equilibrium to be feasible (Fig. 10), and the level of cooperation (x_{CC}^*) in this cooperative equilibrium (Fig. 11):

21

$$\alpha_{Min} = \frac{1}{1 - \sqrt{rp} - \sqrt{(1-r)(1-p)}}$$

$$x^*_{CC} = \frac{m^2 + \delta(1-2p)\big(\delta - 2(1-r)\big) + |m - \delta(1-2p)|\sqrt{(m-\delta)^2 - 4rp\delta}}{2\delta(m^2(1-\delta) + \delta(r-p)^2)}$$

where $m \equiv 1 - r - p$.

[Figure 10]

[figure 11]

In accord with intuition, we find that the normalized payoffs r and p have a polar opposite effect on the behaviour of the system. Thus, if either the normalized "reward payoff" r *increases* or the normalized "punishment payoff" p *falls*, there is an expansion of the region $\{\alpha: \alpha \geq \alpha_{Min}\}$ where a cooperative equilibrium arises, and the level of cooperation achieved in that equilibrium increases.

Conclusions

Some type of conditional behaviour is obviously required for the evolution of cooperation in contexts where individuals interact in an "unstructured environment" with fully random (re)matching. To gain a stark and systematic understanding of the problem, we have focused on a simple model with an unbiased set of possible strategies, where all individuals share the same capacity to react to the most recent action of their partner. In this context, conditional dissociation has been shown to play the key role in the emergence of a (partially) cooperative regime in which, as typically observed in nature, cooperators coexist with defectors (Doebeli et al., 2004). The key factor underlying the high level and eventual predominance of cooperation is the expected

22

lifespan of the individuals. For conditional dissociation to be effective in the emergence of cooperation, it is sufficient that lifespans are only moderately long. Under these conditions, our results starkly show that, in the long run, the population becomes mostly composed of cooperating individuals who leave their partner as soon as the latter defects, i.e. who apply a simple rule of conditional dissociation. We have also obtained a general closed-form approximation for the level of cooperation that shows how this level varies with the expected lifespan of individuals and with the payoffs of the game.

The research reported here lends itself to multiple extensions. An immediate one, for example, would be to allow for the possibility that individuals condition their decision on the last outcome – this would allow, specifically, for strategies such as Pavlov (Nowak and Sigmund, 1993). Other generalizations would include costs of leaving (see supplementary material), or the option to avoid interaction altogether and thus ensure some safe payoff (Aktipis, 2004; Hauert et al., 2007; Janssen, 2008; Mathew and Boyd, 2009; Tesfatsion, 1997). Our preliminary explorations of these extended frameworks suggest that a cooperative regime continues to arise for long enough lifespans, still building upon conditional dissociation as the main mechanism at work.

Acknowledgements

Reference List

Aktipis, C. A., 2004. Know when to walk away: contingent movement and the evolution of cooperation. J. Theor. Biol. 231, 249-260., doi:10.1016/j.jtbi.2004.06.020.

Aktipis, C. A., 2006. Recognition memory and the evolution of cooperation: How simple strategies succeed in an agent-based world. Adapt. Behav. 14, 239-247., doi:10.1177/105971230601400301.

Ashlock, D., Smucker, M. D., Stanley, E. A., Tesfatsion, L., 1996. Preferential partner selection in an evolutionary study of Prisoner's Dilemma. Biosystems 37, 99-125., doi:10.1016/0303-2647(95)01548-5.

Axelrod, R. 1984. The Evolution of Cooperation. Basic Books, New York.

Axelrod, R., 2000. On Six Advances in Cooperation Theory. Analyse & Kritik 22, 130-151.

Benaim, M., Weibull, J. W., 2003. Deterministic approximation of stochastic evolution in games. Econometrica 71, 873-903.

Bendor, J., Swistak, P., 1995. Types of evolutionary stability and the problem of cooperation. Proc. Natl. Acad. Sci. U. S. A. 92, 3596-3600.

Binmore, K., Samuelson, L., Vaughn, R., 1995. Musical chairs: Modelling noisy evolution. Games Econ. Behav. 11, 1-35., doi:10.1006/game.1995.1039.

Björnerstedt, J., Weibull, J. 1996. Nash equilibrium and evolution by imitation. In: K.Arrow et al. (Ed.), The Rational Foundations of Economic Behavior, Macmillan, pp. 155-181.

Coricelli, G., Fehr, D., Fellner, G., 2004. Partner Selection in Public Goods Experiments. J. Conflict. Resolut. 48, 356-378., doi:10.1177/0022002704264143.

de Vos, H., Smaniotto, R., Elsas, D. A., 2001. Reciprocal altruism under conditions of partner selection. Ration. Soc. 13, 139-183., doi:10.1177/104346301013002001.

Doebeli, M., Hauert, C., 2005. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. Ecology Letters 8, 748-766., doi:10.1111/j.1461-0248.2005.00773.x.

Doebeli, M., Hauert, C., Killingback, T., 2004. The Evolutionary Origin of Cooperators and Defectors. Science 306, 859-862., doi:10.1126/science.1101456.

24

Edmonds, B., Hales, D., 2003. Replication, replication and replication: Some hard lessons from model alignment. J. Artif. Soc. Soc. Simulat. 6.

Eshel, I., Cavalli-Sforza, L. L., 1982. Assortment of encounters and evolution of cooperativeness. Proc. Natl. Acad. Sci. U. S. A. 79, 1331-1335.

Fletcher, J. A., Doebeli, M., 2006. How altruism evolves: assortment and synergy. J. Evol. Biol. 19, 1389-1393., doi:10.1111/j.1420-9101.2006.01146.x.

Fletcher, J. A., Doebeli, M., 2009. A simple and general explanation for the evolution of altruism. Proc. Biol. Sci. 276, 13-19., doi:10.1098/rspb.2008.0829.

Fletcher, J. A., Zwick, M., 2004. Strong altruism can evolve in randomly formed groups. J. Theor. Biol. 228, 303-313., doi:10.1016/j.jtbi.2004.01.004.

Fujiwara-Greve, T., Okuno-Fujiwara, M., 2009. Voluntarily Separable Repeated Prisoner's Dilemma. Rev Econ Stud 76, 993-1021., doi: 10.1111/j.1467-937X.2009.00539.x

Hales, D., Edmonds, B., 2003. Evolving Social Rationality for MAS using "Tags". Proceedings of the International Conference on Autonomous Agents 2, 497-503.

Hamilton, I. M., Taborsky, M., 2005. Contingent movement and cooperation evolve under generalized reciprocity. Proc. Biol. Sci. 272, 2259-2267., doi:10.1098/rspb.2005.3248.

Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002a. Replicator Dynamics for Optional Public Good Games. J. Theor. Biol. 218, 187-194., doi:10.1006/jtbi.2002.3067.

Hauert, C., De Monte, S., Hofbauer, J., Sigmund, K., 2002b. Volunteering as Red Queen mechanism for cooperation in public goods games. Science 296, 1129-1132., doi:10.1126/science.1070582.

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., Sigmund, K., 2007. Via freedom to coercion: The emergence of costly punishment. Science 316, 1905-1907., doi:10.1126/science.1141588.

Hauk, E., 2003. Multiple Prisoner's Dilemma Games with(out) an Outside Option: an Experimental Study. Theor. Decis. 54, 207-229., doi:10.1023/A:1027385819400.

Hauk, E., Nagel, R., 2001. Choice of Partners in Multiple Two-Person Prisoner's Dilemma Games: An Experimental Study. J. Conflict. Resolut. 45, 770-793., doi:10.1177/0022002701045006004.

Hayashi, N., 1993. From Tit-for-Tat to Out-for-Tat: The Dilemma of the Prisoner's Network. Riron to Hoho/Sociological Theory and Methods 8, 19-32.

Hruschka, D. J., Henrich, J., 2006. Friendship, cliquishness, and the emergence of cooperation. J. Theor. Biol. 239, 1-15., doi:10.1016/j.jtbi.2005.07.006.

25

Imhof, L. A., Nowak, M. A., 2010. Stochastic evolutionary dynamics of direct reciprocity. Proc. Biol. Sci. 277, 463-468., doi:10.1098/rspb.2009.1171.

Jansen, V. A. A., van Baalen, M., 2006. Altruism through beard chromodynamics. Nature 440, 663-666., doi:10.1038/nature04387.

Janssen, M. A., 2008. Evolution of cooperation in a one-shot Prisoner's Dilemma based on recognition of trustworthy and untrustworthy agents. J. Econ. Behav. Organ. 65, 458-471., doi:10.1016/j.jebo.2006.02.004.

Jin, N., Hayashi, N., Shinotsuka, H., 1996. An experimental study of prisoner's dilemma network: The formation of commitment among selective dyads. Japanese Journal of Experimental Social Psychology 33, 21-30.

Joyce, D., Kennison, J., Densmore, O., Guerin, S., Barr, S., Charles, E., Thompson, N. S., 2006. My Way or the Highway: a More Naturalistic Model of Altruism Tested in an Iterative Prisoners' Dilemma. J. Artif. Soc. Soc. Simulat. 9.

Le, S., Boyd, R., 2007. Evolutionary dynamics of the continuous iterated Prisoner's dilemma. J. Theor. Biol. 245, 258-267., doi:10.1016/j.jtbi.2006.09.016.

Macy, M., Skvoretz, J., 1998. The evolution of trust and cooperation between strangers: a computational model. Am. Socio. Rev. 63, 638-660.

Mathew, S., Boyd, R., 2009. When does optional participation allow the evolution of cooperation? Proc. Biol. Sci. 276, 1167-1174., doi:10.1098/rspb.2008.1623.

McDonald, A., Sen, S., 2006. The success and failure of tag-mediated evolution of cooperation. Lect. Notes Comput. Sci. 3898, 155-164., doi:10.1007/11691839_9.

McNamara, J. M., Barta, Z., Fromhage, L., Houston, A. I., 2008. The coevolution of choosiness and cooperation. Nature 451, 189-192., doi:10.1038/nature06455.

Németh, A., Takács, K., 2010. The paradox of cooperation benefits. J. Theor. Biol. In Press, Corrected Proof., doi:10.1016/j.jtbi.2010.02.005.

Nowak, M. A., 2006. Five rules for the evolution of cooperation. Science 314, 1560-1563., doi:10.1126/science.1133755.

Nowak, M., Sigmund, K., 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature 364, 56-58., doi:10.1038/364056a0.

Nowak, M. A., Tarnita, C. E., Antal, T., 2010. Evolutionary dynamics in structured populations. Phil. Trans. R. Soc. B 365, 19-30., doi:10.1098/rstb.2009.0215.

Orbell, J., Dawes, R. M., 1991. A "Cognitive Miser" Theory of Cooperators' Advantage. Am. Polit. Sci. Rev. 85, 515-528.

Orbell, J., Dawes, R. M., 1993. Social welfare, cooperators' advantage, and the option of not playing the game. Am. Socio. Rev. 58, 787-800.

Pacheco, J. M., Traulsen, A., Nowak, M. A., 2006. Active linking in evolutionary games. J. Theor. Biol. 243, 437-443., doi:10.1016/j.jtbi.2006.06.027.

Pennisi, E., 2005. How Did Cooperative Behavior Evolve? Science 309, 93., doi:10.1126/science.309.5731.93.

Riolo, R. L., Cohen, M. D., Axelrod.R., 2001. Evolution of cooperation without reciprocity. Nature 414, 441-443., doi:10.1038/35106555.

Rowley, I. 1983. Re-mating in birds. In: P.Bateson (Ed.), Mate Choice, Cambridge University Press, pp. 331-360.

Sachs, J. L., 2006. Cooperation within and among species. J. Evol. Biol. 19, 1415-1418., doi:10.1111/j.1420-9101.2006.01152.x.

Santos, F. C., Pacheco, J. M., Lenaerts, T., 2006. Cooperation Prevails When Individuals Adjust Their Social Ties. PLoS Comput Biol 2, e140., doi:10.1371/journal.pcbi.0020140.

Schuessler, R., 1989. Exit Threats and Cooperation under Anonymity. J. Conflict. Resolut. 33, 728-749., doi:10.1177/0022002789033004007.

Sherratt, T. N., Roberts, G., 1998. The evolution of generosity and choosiness in cooperative exchanges. J. Theor. Biol. 193, 167-177., doi:10.1006/jtbi.1998.0703.

Simms, E. L., Taylor, D. L., Povich, J., Shefferson, R. P., Sachs, J. L., Urbina, M., Tausczik, Y., 2006. An empirical test of partner choice mechanisms in a wild legume-rhizobium interaction. Proc. Biol. Sci. 273, 77-81., doi:10.1098/rspb.2005.3292.

St-Pierre, A., Larose, K., Dubois, F., 2009. Long-term social bonds promote cooperation in the iterated Prisoner's Dilemma. Proc. Biol. Sci. 276, 4223-4228., doi:10.1098/rspb.2009.1156.

Szabó, G., Hauert, C., 2002. Evolutionary prisoner's dilemma games with voluntary participation. Phys. Rev. E 66, 062903., doi:10.1103/PhysRevE.66.062903.

Tanimoto, J., 2007. Does a tag system effectively support emerging cooperation? J. Theor. Biol. 247, 756-764., doi:10.1016/j.jtbi.2007.03.033.

Taylor, P. D., Day, T., 2004. Behavioural evolution: Cooperate with thy neighbour? Nature 428, 611-612., doi:10.1038/428611a.

Tesfatsion, L. 1997. A trade network game with endogenous partner selection. In: H.Amman, B.Rustem, A.B.Whinston (Eds.), Computational Approaches to Economic Problems, Kluwer Academic Publishers, pp. 249-269.

The EdK-Group, 2000. Exit, Anonymity and the Chances of Egoistical Cooperation. Analyse & Kritik 22, 114-129.

Thomas, E. A. C., Feldman, M. W., 1988. Behavior-Dependent Contexts for Repeated Plays of the Prisoner's Dilemma. J. Conflict. Resolut. 32, 699-726., doi:10.1177/0022002788032004005.

Vanberg, V. J., Congleton, R. D., 1992. Rationality, Morality, and Exit. Am. Polit. Sci. Rev. 86, 418-431.

Yamagishi, M., Hayashi, N., Jin, N. 1994.

Social Dilemmas and Cooperation, , pp. 233-250.

Figure Captions

Fig. 1. Sequence of events within each time period. The "Remaining Pairs" and "Singles" at the end of a period are identical to the "Existing Pairs" and "Singles" in the next period.

Fig. 2. Evolution of the percentage of DD outcomes for 18 different runs, each run starting with the whole population using one of the 18 different strategies. Parameterization: $N = 1000$, $\alpha = 5$, $\mu = 0.05$, $T = 4$, $R = 3$, $P = 1$, $S = 0$.

Fig. 3. Evolution of the percentage of CC outcomes for 18 different runs, each run starting with the whole population using one of the 18 different strategies. Parameterization: $N = 1000$, $\alpha = 100$, $\mu = 0.05$, $T = 4$, $R = 3$, $P = 1$, $S = 0$.

Fig. 4. Evolution of the percentage of CC outcomes for 18 different runs, each run starting with the whole population using one of the 18 different strategies. Parameterization: $N = 1000$, $\alpha = 20$, $\mu = 0.05$, $T = 4$, $R = 3$, $P = 1$, $S = 0$.

Fig. 5. Fraction of periods spent in the cooperative and in the non-cooperative regimes as a function of the expected life α. The values in each column are compiled over 10^3 simulation runs. Every run measured between periods $3 \cdot 10^3$ and 10^4, with random initial conditions. Parameterization: $N = 1000$, $\mu = 0.05$, $T = 4$, $R = 3$, $P = 1$, $S = 0$.

Fig. 6. Composition of strategies in the cooperative regime computed in Fig. 5.

Fig. 7. Average values of the level of cooperation (x_{CC}) and of the proportion of Tit-For-Tat strategists (% TFT) as a function of individuals' expected life, in a model where the option to leave has been removed by considering only the 8 strategies that always stay

29

with their partner. The values for each point are compiled over 10^3 simulation runs. Every run measured between periods $3 \cdot 10^3$ and 10^4, with random initial conditions. Parameterization: $N = 1000$, $T = 4$, $R = 3$, $P = 1$, $S = 0$.

Fig. 8. Trajectories and critical points of the mean field for the reduced system with (a) $\alpha = 1$, (b) $\alpha = 10$ and (c) $\alpha = 20$. Parameterization: $T = 4$, $R = 3$, $P = 1$, $S = 0$, $\mu = 0$. Asymptotically stable critical points are shown in yellow.

Fig. 9. Average values of the level of cooperation (x_{CC}) as a function of the individuals' expected life, both in the cooperative equilibrium of the mean-field approximation of the reduced system (MF) and in the cooperative regime of the stochastic simulations of the non-reduced system ($\hat{x}_{CC}^{CR}$), for different values of μ. Parameterization: $N = 1000$, $T = 4$, $R = 3$, $P = 1$, $S = 0$.

Fig. 10. Minimum value of the expected life α (α_{Min}) for the existence of the cooperative equilibrium, as a function of the normalized payoffs for reward (r) and for punishment (p).

Fig. 11. Level of cooperation x_{CC} for $\alpha = 20$ in the critical point corresponding to the cooperative equilibrium, as a function of the normalized payoffs for reward (r) and for punishment (p).
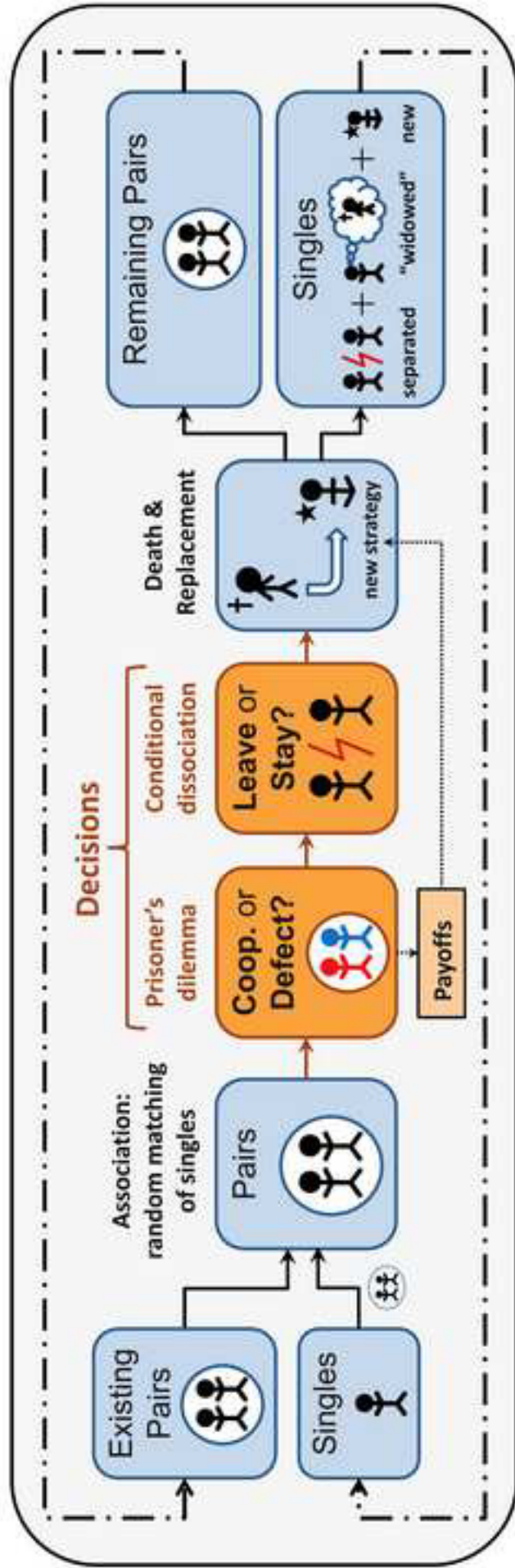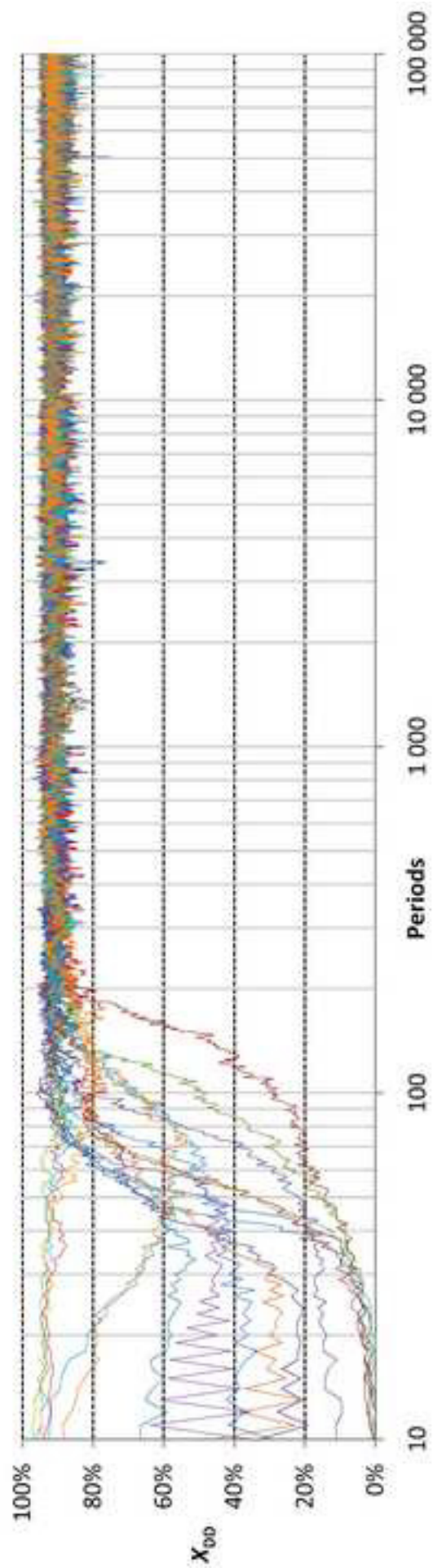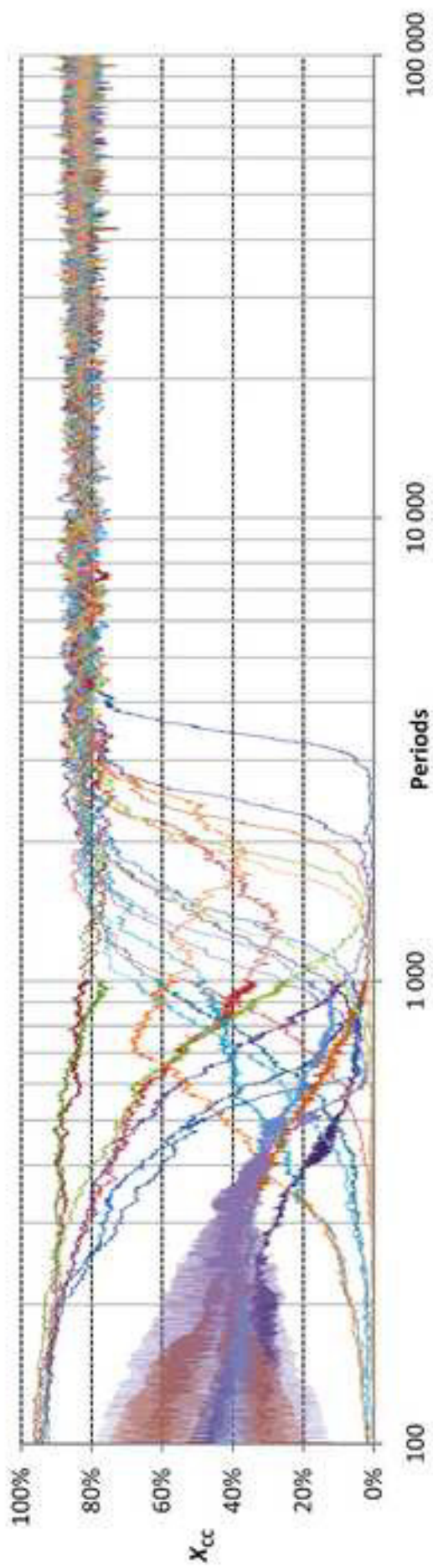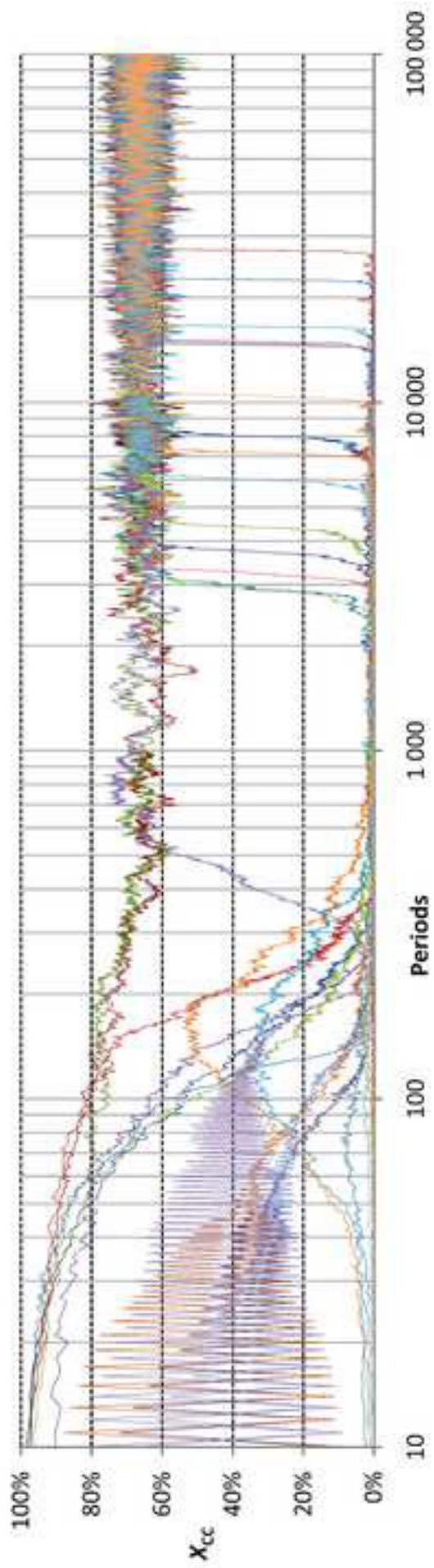
Figure 1

Figure 2

Figure 3

Figure 4
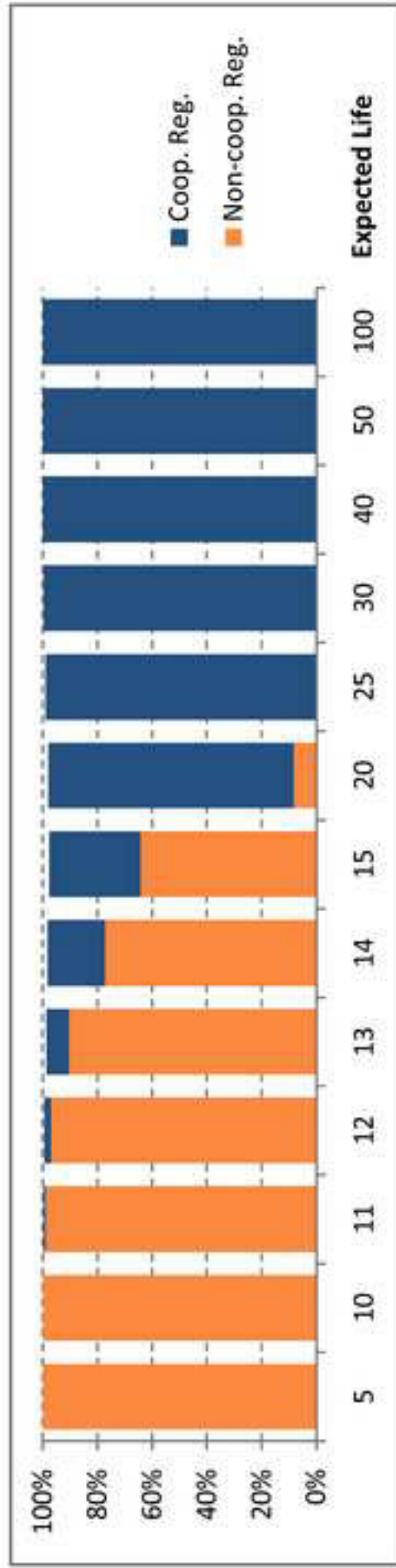
Figure 5

Figure 6

4. Figure 7

4. Figure 8

4. Figure 9

Figure 9

4. Figure 11