COOPERATION, IMITATION AND CORRELATED MATCHING

Javier Rivas, University of Leicester, UK

Working Paper No. 09/12
June 2009
Cooperation, Imitation and Correlated Matching

JAVIER RIVAS

University of Leicester

June 10, 2009

Abstract

We study a setting where players are matched into pairs to play a Prisoners’ Dilemma game. In the model presented, players are not rational; they simply imitate the more successful actions they observe. Furthermore, a certain correlation is added to the matching process: players that belong to a pair were both parties cooperate repeat partner next period whilst all other players are randomly matched into pairs. Although cooperation vanishes for any initial interior condition under complete random matching, the correlation in the matching process considered in this paper makes a significant amount of cooperation the unique outcome under mild conditions. Furthermore, it is shown that no separating equilibrium, i.e. a situation where cooperators and defectors are not matched together, exits.

JEL Classification: C71, C73, C78.

Keywords: Cooperation, Correlated Matching, Imitation, Prisoners’ Dilemma.

*I would like to thank Subir Bose, Gianni de Fraja, Friederike Mengel, Mark Le Quement, Carlos Oyazun and Ludovic Renou for comments and suggestions. I would also like to thank the attendants to the Internal Seminar series at the University of Leicester for useful comments.

†javier.rivas@le.ac.uk. Department of Economics, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom. www.le.ac.uk/users/jr168/index.htm.
1 Introduction

Individuals learn by imitation when their choices are based on the alternatives they observe others choose. Under mild conditions, a population whose behavior follows imitation learns not to play dominated actions (see, for instance, Schlag (1998) Remark 6). Thus, if every period imitative players are randomly matched to play a Prisoner’s Dilemma game, cooperation vanishes in the long run. Given the importance of cooperation and its constant presence in societies (see Axelrod (1984)), and the relevance of imitation for modeling bounded rational behavior (see, for example, Banerjee (1992), Eshel et al (1998) or Ellison and Fudenberg (1995)), the question we raise is: can cooperation survive when players learn by imitation?

We answer this question by exploring the mechanism by which players are matched to play a Prisoners’ Dilemma game. The novelty of this paper is that a certain correlation is introduced in the matching process: players who cooperated with each other last period meet again in the next period whilst the rest of players are randomly matched into pairs. This matching mechanism captures the simple idea that a player should have no incentives to repeat partner unless the partner played cooperatively last period.

The matching process considered in this paper adds a positive externality to playing cooperatively: in a situation where two players cooperate, switching action has the disadvantage that next period a new opponent, who might not be so keen on playing cooperatively, is faced. However, this argument is rational. Does the previous statement holds when players simply imitate each other? The answer is positive. To understand this consider the simple situation where the population is divided into two groups: players who cooperate and are matched with another player also cooperating and players who do not cooperate and are matched with another player not cooperating. The way matching works implies that pairs where both players cooperate repeat partner. Therefore, in this case, players that cooperate enjoy more payoff than those not cooperating. Hence, non-cooperative players may imitate cooperative ones, making the survival of cooperation possible.

In the results of this paper three main conclusions are achieved: First, under mild conditions and for any interior initial condition, the survival of a non-negligible amount of cooperation is guaranteed. That is, the situation where no player cooperates is not stable if mild conditions on the payoff matrix and/or the specific imitative rule employed are satisfied. Second, no separating equilibrium exists. This means that, apart from the equilibria on the boundaries, a situation where cooperative players do not face non-cooperative ones is not an equilibrium. Finally, we show that under certain conditions all players cooperating is an stable outcome.

We believe that a value added to this paper lies in the way some of our results are derived.
For dealing with the high-order system of difference equations determining the evolution of the population, we employ what is known as the continuous time approximation (Benaim and Weibull (2003)). Although this technique is usually applied to eliminate the randomness of a certain model, we apply it here to decrease the order of the difference equations as well as to move from discrete time to continuous time. We prove that both the original model and the approximated model behave similarly when the system is close to the end-points, i.e. all cooperate, no one cooperates. Thus, results obtained from the analysis of the approximated model can be extrapolated to the original model.

To our knowledge, only Levine and Pesendorfer (2007) and Bergstrom (2003) study similar settings to the one considered in this paper. Levine and Pesendorfer (2007) show that cooperation can survive within a population who learns by imitation if each player holds some information about the strategy of the player with whom she is matched. Bergstrom (2003) proves conditions under which cooperation survives in an evolutionary model where players are either cooperators or defectors, and are more likely to face a player of their same type.

The difference between this paper and Levine and Pesendorfer (2007) lies in that in our model there is a set of matches that are anonymous whilst in Levine and Pesendorfer (2007) all matches are non-anonymous to a certain degree. The present paper differs from Bergstrom (2003) in that players can change their actions from one period to another. Thus, playing cooperatively in the present period is no guarantee of exhibiting a cooperative behavior in the next period.

The issue of partner selection in cooperative games has recently attracted attention from experimental economists. Duffy and Ochs (2009) conduct an experiment where a Prisoners Dilemma game with two treatments is considered. In the first treatment, matching is completely random whereas in the second one each player always repeats partner. The authors find that cooperation does not emerge in the random matching setting while it does in the fixed pairs treatment. Yang et al (2007) present an experiment where a Prisoner Dilemma game is played and individuals with similar histories are more likely to be matched together. Their results show that cooperation has a higher chance of survival when a history-dependent correlation is added to the matching process. Grimm and Mengel (2009) develop an experiment where players choose between two Prisoner’s Dilemma games that differ in the gains from defection. Choosing the game with lower gains signals the player’s willingness to cooperate. Grimm and Mengel find that this self selection significantly increases the amount of cooperation.

The rest of the paper is organized as follows. In Section 2, we develop the model. Section 3 presents the main analysis and the results. In Section 4, we discuss the robustness of our findings and assumptions as well as present some extensions. Finally, Section 5 concludes.
2 The Model

Consider a continuum of identical players uniformly distributed on the interval \([0, 1]\) with the standard Borel-Lebesgue measure. At the beginning of each period \(t = 0, 1, 2, \ldots\), every player is paired with another one and plays the following symmetric stage game against her partner:

<table>
<thead>
<tr>
<th></th>
<th>(C)</th>
<th>(D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C)</td>
<td>(a, a)</td>
<td>(b, c)</td>
</tr>
<tr>
<td>(D)</td>
<td>(c, b)</td>
<td>(d, d)</td>
</tr>
</tbody>
</table>

Where \(C\) stands for cooperate and \(D\) stands for defect. The stage game above has the standard Prisoners' Dilemma structure: \(c > a > d > b\) with \(a, b, c, d \in \mathbb{R}\).

After the stage game is played, all pairs where at least one player chose \(D\) are broken while the rest of pairs are maintained. After that, unpaired players are randomly matched into pairs. The distribution of pairs at the beginning of period \(t = 0\) is given.

Given the description above, at the beginning of each period \(t \geq 1\) the population is divided into three sets: players who played \(C\) last period and faced an opponent who also played \(C\), \(\gamma\), players who played \(C\) but faced an opponent who played \(D\), \(\sigma\), and the rest, denoted by \(1 - \gamma - \sigma\). We use \(\gamma\), \(\sigma\) and \(1 - \sigma - \gamma\) to denote exchangeably both the sets and their respective measure. Thus, for instance, \(\gamma\) is both the set of players who played \(C\) and faced an opponent who also chose \(C\), and the measure (fraction) of players who played \(C\) and faced an opponent who also chose \(C\). Therefore, we have that \(\gamma \in [0, 1]\), \(\sigma \in [0, 1]\) and \(1 - \gamma - \sigma \in [0, 1]\). Evidently, \(\sigma + \gamma \leq 1\) with equality only in the case when \(\gamma = 1\) (if \(\sigma + \gamma = 1\) then all players chose \(C\) and, thus, all players faced another one playing \(C\)). Notice that the fraction of players who maintain partner equals \(\gamma\). Furthermore, note that all players in \(\sigma\) are matched with a player in \(1 - \gamma - \sigma\) and, thus, \(\sigma \leq 1 - \gamma - \sigma\). Finally, define \(\Omega\) as

\[
\Omega = \{(\gamma, \sigma) \in \mathbb{R}^2_+ : \gamma + \sigma < 1 \cup (\gamma, \sigma) = (1, 0)\}.
\]

Whenever we refer to interior points we mean \((\gamma, \sigma) \in \text{int}(\Omega)\).

Players follow very simple decision rules. In particular, they observe the action and payoff of a random individual and base their choice of action for the stage game on this information plus the information from own action and payoff. All players in the population are equally likely to be observed. Since we are dealing with a continuous population, results presented in this paper do not depend on how many players are observed.
Let \( A \in \{C, D\} \) be the action set and let \( P(\{i, a_i, \pi_i\}\{j, a_j, \pi_j\}) \in [0, 1] \) be the probability with which player \( i \in [0, 1] \) changes action if she, who played action \( a_i \in A \) and obtained payoff \( \pi_i \in \mathbb{R} \), observes player \( j \in [0, 1] \), who chose action \( a_j \in A \) and achieved payoff \( \pi_j \in \mathbb{R} \). Some assumptions on \( P \) are needed for the analysis:

**Assumptions.**

1. If \( a_i = a_j \) then \( P(\{i, a_i, \pi_i\}\{j, a_j, \pi_j\}) = 0 \),
2. \( P(\{i, a_i, \pi_i\}\{j, a_j, \pi_j\}) > 0 \) if and only if \( \pi_i < \pi_j \) and,
3. for all \( i, j \in [0, 1] \) and all \( a_i, a_j \in A \):
   - if \( \pi_j > \pi_j^* \) then \( P(\{i, a_i, \pi_i\}\{j, a_j, \pi_j\}) \geq P(\{i, a_i, \pi_i^*\}\{j, a_j, \pi_j^*\}) \),
   - if \( \pi_i < \pi_i^* \) then \( P(\{i, a_i, \pi_i\}\{j, a_j, \pi_j\}) \geq P(\{i, a_i, \pi_i^*\}\{j, a_j, \pi_j^*\}) \).

The first two assumptions are standard in imitation models (see, for instance, Schlag (1998)). Assumption 1 implies that players change their action only if the player they observe played a different action than the one they chose. Assumption 2 means that there is a positive probability of changing action if and only if observed action yielded more payoff than own action. The third assumption is a monotonicity condition that relates to reinforcement learning models (see, for example, Börgers et al (2004) and Rustichini (1999)). It means that the probability of changing action is weakly increasing in observed payoff and weakly decreasing in own payoff.

We simplify notation when using the function \( P(\{i, a_i, \pi_i\}\{j, a_j, \pi_j\}) \) as follows: Denote by \( P_D^\gamma : A^2 \times \mathbb{R}^2 \to [0, 1] \) the probability with which a player in \( \gamma \) changes to \( D \). Let \( P_D^\sigma : A^2 \times \mathbb{R}^2 \to [0, 1] \) be the probability with which a player who belongs to \( \sigma \) changes to \( D \). Finally, denote by \( P_C^{1-\gamma-\sigma} : A^2 \times \mathbb{R}^2 \to [0, 1] \) the probability with which a player in \( 1-\gamma-\sigma \) changes to \( C \).

Assumptions 1–3 impose some restrictions on the functional forms of \( P_D^\gamma, P_D^\sigma \) and \( P_C^{1-\gamma-\sigma} \). The function \( P_D^\gamma \) is only positive if the player in \( \gamma \) observes a player in \( 1-\gamma-\sigma \) (Assumption 1) who faced a player in \( \sigma \) (Assumption 2). In this case, the payoff of observed player equals \( c \) while own payoff equals \( a \). Thus, we can write \( P_D^\gamma \) for \( \gamma > 0 \) as

\[
P_D^\gamma = (1 - \gamma - \sigma) \frac{\sigma}{1 - \gamma} f(c,a)
\]

for some function \( f : \mathbb{R}^2 \to [0, 1] \). The two arguments in \( f \) are observed payoff and own payoff respectively. The function \( f \) is weakly increasing in its first argument and weakly decreasing in its second argument by Assumption 3. Furthermore, by Assumption 2, \( f(\pi', \pi) = 0 \) for
any \( \pi' > \pi \). If \( \gamma = 0 \) then \( \sigma = 0 \) and \( P_0^D \) is well defined and continuously differentiable for all \((\gamma, \sigma) \in \Omega\).

The function \( P_0^D \) is only positive if the player in \( \sigma \) observes a player in \( 1 - \gamma - \sigma \). In this case, two different situations arise: If the player observed faced a player in \( \sigma \), then observed payoff equals \( c \) and own payoff equals \( b \). On the other hand, if the observed player faced an opponent in \( 1 - \gamma - \sigma \), then observed payoff equals \( d \) and own payoff equals \( b \). Therefore, we have that

\[
P_0^\sigma = (1 - \gamma - \sigma) \left[ \frac{\sigma}{1 - \gamma} f(c, b) + \frac{1 - \sigma - \gamma}{1 - \gamma} f(d, b) \right]
\]

if \( \gamma > 0 \), \( P_0^\sigma = 0 \) otherwise.

Finally, \( P_C^{1-\gamma-\sigma} \) is only positive if the player in \( 1 - \gamma - \sigma \) faced a player also in \( 1 - \gamma - \sigma \) and observed player that belongs to \( \gamma \). In this case, observed payoff equals \( a \) while own payoff equals \( d \). Hence, we have that

\[
P_C^{1-\gamma-\sigma} = \frac{(1 - \gamma - \sigma)}{1 - \gamma} f(a, d)
\]

if \( \gamma > 0 \), \( P_C^{1-\gamma-\sigma} \) otherwise.

Let \( \gamma^t \) and \( \sigma^t \) denote the values of \( \gamma \) and \( \sigma \) respectively at each point in time \( t = 0, 1, 2, \ldots \) before the stage game is played with \((\gamma^0, \sigma^0) \in \Omega \) given. At \( t = 0 \) and prior to the starting of the game, all players not in \( \gamma^0 \) are randomly and uniformly matched into pairs. For notational convenience the argument \( t \) in the functions \( P_0^\sigma, P_D^\sigma \) and \( P_C^{1-\gamma-\sigma} \) is omitted.

The model just described can be expressed with the following system of difference equations:

\[
s^{t+1} = \gamma^t (1 - P_0^D) + \sigma^t (1 - P_D^0) + (1 - \gamma^t - \sigma^t) P_C^{1-\gamma-\sigma} - \gamma^t (1 - P_0^D)^2 - \left( \sigma^t (1 - P_D^0) + (1 - \gamma^t - \sigma^t) P_C^{1-\gamma-\sigma} \right)^2 \]

\[
\gamma^{t+1} = \gamma^t (1 - P_0^D)^2 + \left( \sigma^t (1 - P_D^0) + (1 - \gamma^t - \sigma^t) P_C^{1-\gamma-\sigma} \right)^2
\]

if \( \gamma^t > 0 \), \( \sigma^{t+1} = \gamma^{t+1} = 0 \) otherwise. Note that if \( t \geq 1 \) then \( \gamma^t = 0 \) implies \( \sigma^t = 0 \). Thus, \( \sigma^{t+1} \) and \( \gamma^{t+1} \) are both well defined and continuously differentiable in all \( \Omega \).

Equation (4) tells us the measure of players who played \( C \) in period \( t \) and faced a player who chose \( D \) in \( t \). The value of \( \sigma^{t+1} \) is computed as follows: The first three terms represent all players who played \( C \) in \( t \) (note that players in \( \gamma^t \) and \( \sigma^t \) played \( C \) in \( t - 1 \) but may have
played \( D \) in \( t \). The fourth term subtracts the pairs in \( \gamma^t \) where both players played \( C \) again in \( t \). Finally, the fifth term subtracts the players not in \( \gamma^t \) who chose \( C \) in \( t \) and faced a player one who also chose \( C \) in \( t \).

Equation (5) is the measure of players who chose \( C \) in period \( t \) and faced an opponent playing \( C \) in \( t \). The value of \( \gamma^{t+1} \) is determined as follows: The first term adds the pairs in \( \gamma^t \) where both players played \( C \) in \( t \). The second term adds the players not in \( \gamma^t \) who choose \( C \) in \( t \) and faced a player who also chose \( C \) in \( t \).

Next, we define what an equilibrium of the model at hands is. Intuitively, an equilibrium is a situation where the measure of players belonging to each of the three sets \( \gamma \), \( \sigma \) and \( 1 - \gamma - \sigma \) does not change. Formally:

**Definition 1.** An equilibrium is a point \((\gamma, \sigma) \in \Omega \) such that \( \gamma^{t+1} = \gamma^t \) and \( \sigma^{t+1} = \sigma^t \) whenever \( \gamma^t = \gamma \) and \( \sigma^t = \sigma \).

Among all equilibria it is useful to single out the separating equilibria. A separating equilibrium is an equilibrium where a fraction of the population play \( C \) against themselves while all other players choose \( D \). That is, in a separating equilibrium, \( \sigma = 0 \), and the population is completely separated between cooperators and defectors.

**Definition 2.** A separating equilibrium is an equilibrium where \( \gamma \in (0, 1) \) and \( \sigma = 0 \).

Further to the definitions of equilibrium, it is necessary for the analysis to distinguish between the different notions of stability. The following definitions are based on Khalil (1995).

**Definition 3.** Let \( B_r(\gamma, \sigma) \) be the ball of radius \( r > 0 \) around the point \((\gamma, \sigma) \in \Omega \). The equilibrium \((\gamma, \sigma) \in \Omega \) is

- **stable** if for any \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that if \((\gamma^0, \sigma^0) \in \Omega \cap B_\delta(\gamma, \sigma) \) then \((\gamma^t, \sigma^t) \in \Omega \cap B_\varepsilon(\gamma, \sigma) \) for all \( t \geq 0 \),
- **unstable** if it is not stable,
- **asymptotically stable** if it is stable and \( \delta > 0 \) can be chosen such that for any \( \kappa < \varepsilon \) if \((\gamma^0, \sigma^0) \in \Omega \cap B_\delta(\gamma, \sigma) \) then
  \[
  \| \lim_{t \to \infty} (\gamma^t, \sigma^t) - (\gamma, \sigma) \| < \kappa,
  \]
  - **a repeller** if there exists a \( \delta > 0 \) such that if \((\gamma^0, \sigma^0) \in \Omega \cap B_\varepsilon(\gamma, \sigma) \) for all \( \varepsilon \in (0, \delta) \) then \((\gamma^t, \sigma^t) \not\in \Omega \cap B_\delta(\gamma, \sigma) \) for some \( t \geq 0 \),
Whenever results from simulations are presented the specific imitative behavior players use follows what is known as the Proportional Imitation Rule (hereafter PIR) due to Schlag (1998). According to the PIR, a action is adapted with a probability equal to the normalized difference between observed and own payoff. Given that $c$ is the maximum payoff achievable and $b$ is the minimum one, the function $f$ for the PIR is given by

$$f(\pi', \pi) = \frac{1}{c - b} (\pi' - \pi)$$

when observed action is different from own action and observed payoff, $\pi'$, is higher than own payoff, $\pi$. If $\pi' \leq \pi$ or observed action is equal to own action, then $f(\pi', \pi) = 0$.

In order to illustrate the behavior of the model, Figure 1 shows the result of a simulation. As it can be observed, during the first periods the amount cooperative players matched with non-cooperative ones, $\sigma$, decreases. This is due to the fact that, during these first stages, most cooperative players enjoy less payoff than cooperative ones. However, as times evolves, more and more cooperative players meet each other. After this grouping stage is over, the payoff from cooperating is on average greater than that from not cooperating. This happens because most cooperative players face players that are also cooperative. The level of cooperation increases from there until all players have adapted the cooperative action.

Figure 1: Simulation: PIR with $a = 0.4$, $b = -0.1$, $c = 0.5$, $d = 0$ and $(\gamma_0, \sigma_0) = (0, 0.5)$
3 Results

3.1 Random Matching

In this subsection we consider the benchmark case of random matching. Under random matching, all pairs are broken after the stage game is played. We show that, under random matching, cooperation vanishes for any interior initial condition. The full analysis of the random matching case is presented in the Appendix; here we restrict our attention to the main result from this analysis.

Proposition 1. Under random matching, for any \((\gamma^0, \sigma^0) \in \text{int}(\Omega)\)

\[
\lim_{t \to \infty} \gamma^t + \sigma^t = 0.
\]

Proof. See Lemma 1 in the Appendix.

As Proposition 1 shows, under random matching cooperation does not survive in the population. This is the known result that under mild monotonicity conditions (Assumption 3) imitation rules out dominated actions (Schlag (1998)).

With random matching, playing cooperatively is always dominated by the non-cooperative behavior. However, once we add a correlation to the matching process, this is no longer the case. If pairs where both players cooperated are preserved, then a positive externality to being cooperative is added. When two players cooperate, they achieve the second highest payoff and, thus, are not too likely to imitate other players. If no player in the pair changes action, they continue to achieve the second highest payoff and, hence, repeated interaction makes cooperation possible.

3.2 Correlated Matching

We now revert back to the case with correlated matching, i.e. pairs were both players cooperated are maintained in the next period. A first result is that there exist no separating equilibria.

Proposition 2. No separating equilibrium exists.

Proof. First, note that from (4) and (5) we have that

\[
(1 - \gamma - \sigma) \gamma_{-\gamma_{-\sigma}} - \gamma P_{\sigma} - \sigma P_{\sigma}^\gamma = 0. \tag{6}
\]
The next step is to show that in an equilibrium with $\sigma = 0$, no pairs in $\gamma$ are ever broken. Assume the contrary, if some pairs are broken that means that some players from $\gamma$ choose $D$. If we are at time $t$, this implies that $\gamma^{t+1} < \gamma^t$ unless some players in $1 - \sigma^t - \gamma^t$ switch to $C$. If this happens, however, we have that some players will be matched against players who chose $D$ in $t$. Hence, if a pair is broken either $\gamma^{t+1} < \gamma^t$ or $\sigma^{t+1} > 0$, a contradiction to the definition of separating equilibrium.

Given that in a separating equilibrium no pairs are ever broken and that $\gamma \in (0,1)$, it follows that all players always choose the same action in the stage game. This implies that players in $\gamma^t$ obtain a payoff of $a$ while players in $1 - \sigma^t - \gamma^t$ obtain a payoff of $d$. Thus, from Assumptions 2 it follows that $P_D^\gamma = 0$ and $P_C^{1-\gamma-\sigma} > 0$. However, when $P_D^\gamma = 0$, equation (6) implies that

$$(1 - \gamma - \sigma) P_C^{1-\gamma-\sigma} = 0.$$  

Since $\gamma \in (0,1)$, $\sigma = 0$ and $P_C^{1-\gamma-\sigma} > 0$, we have that $(1 - \gamma - \sigma) P_C^{1-\gamma-\sigma} > 0$, a contradiction to the condition above. 

The intuition behind the result above is straightforward: In a separating equilibrium, cooperative players, $\gamma$, obtain a payoff of $a$ whilst all the other players, $1 - \gamma$, obtain a payoff of $d < a$. Thus, non-cooperative players imitate cooperative ones but cooperative players do not imitate non-cooperative ones. Therefore, the situation with complete separation between cooperators and defectors is not an equilibrium.

The system of difference equations (4) and (5) is highly nonlinear. In particular, the two expressions on right hand side of both (4) and (5) are polynomials of degree six. For studying the behavior of the model at hands, we employ what is known as the continuous time approximation (see, for instance, Benaim and Weibull (2003)). This approximation consists of taking the time step between interactions to zero while the response to each interaction is also taken to zero and at the same rate. The continuous time approximation is more deeply explained below. Numerical and analytical comparisons between the behavior of the original discrete time model and the behavior of the continuous time approximated model can be found in Section 4.1.

In the continuous time approximation, each time interval is divided into $1/\delta$ subintervals with $\delta \in (0,1]$. Thus, the time scale is then $t = 0, 2\delta, \ldots, 1, 1 + \delta, 1 + 2\delta, \ldots$. Within each of these subintervals, only a fraction $\delta$ of the original response to each interaction occurs. That is, if the probability of changing action in the original model is $P$, then the probability of changing action in the continuous time version is given by $\delta P$. Furthermore, in the continuous time version of the model only a fraction $\delta$ of the couples where both players cooperated
are maintained. Evidently, if $\delta = 1$ the two models are equivalent. The continuous time approximation is obtained by taking the limit when $\delta$ tends to zero.

The usefulness of the continuous time approximation lies in the fact that since $\delta$ is made arbitrarily small, all the terms of order $\delta$ and higher are negligible and, hence, can be ignored.

The continuous time model is given by

$$\sigma^{t+\delta} = \gamma^t (1 - \delta \sigma_D^t) \sigma_D^t + \sigma^t (1 - \delta \sigma_D^t) + (1 - \gamma^t - \sigma^t) \delta \sigma_{C}^{1-\gamma - \sigma}$$

$$\gamma^{t+\delta} = \gamma^t (1 - \delta \sigma_D^t)^2$$

$$+ \delta \frac{(\sigma^t (1 - \delta \sigma_D^t) + (1 - \gamma^t - \sigma^t) \delta \sigma_{C}^{1-\gamma - \sigma})^2}{1 - \gamma^t}.$$ 

Which can be rewritten as

$$\frac{\sigma^{t+\delta} - \sigma^t}{\delta} = \gamma^t \sigma_D^t - \sigma^t \sigma_D^t + (1 - \gamma^t - \sigma^t) \sigma_{C}^{1-\gamma - \sigma} - \frac{(\sigma^t)^2}{1 - \gamma^t} + o(\delta),$$

$$\frac{\gamma^{t+\delta} - \gamma^t}{\delta} = -2\gamma^t \sigma_D^t + \frac{(\sigma^t)^2}{1 - \gamma^t} + o(\delta)$$

where all the terms of order $\delta$ and higher are denoted by $o(\delta)$.

Therefore, if we take $\delta$ to zero, define $\dot{x} = \frac{x^{t+\delta} - x^t}{\delta}$ with $x = \{\gamma, \sigma\}$ and drop the superscript $t$ for notational convenience, we can rewrite the equations determining the evolution of the population as follows:

$$\dot{\sigma} = \gamma \sigma_D - \sigma \sigma_D + (1 - \gamma - \sigma) \sigma_{C}^{1-\gamma - \sigma} - \frac{\sigma^2}{1 - \gamma},$$

$$\dot{\gamma} = -2\gamma \sigma_D + \frac{\sigma^2}{1 - \gamma}.$$

Finally, if we substitute the value of $P_D^\gamma, P_D^\sigma$ and $P_{C}^{1-\gamma - \sigma}$ from equations (1), (2) and (3) into the two equations above we obtain

$$\dot{\sigma} = \frac{1}{1 - \gamma} \left[ (1 - \gamma - \sigma) \gamma \sigma f(c, a) + (1 - \gamma - \sigma) \sigma^2 f(c, b) + (1 - \gamma - \sigma)^2 \sigma f(d, b) + (1 - \gamma - \sigma)^2 \gamma f(a, d) - \sigma^2 \right],$$

$$\dot{\gamma} = \frac{1}{1 - \gamma} \left[ -2(1 - \gamma - \sigma) \gamma \sigma f(c, a) + \sigma^2 \right]$$

if $\gamma > 0$, $\dot{\sigma} = \dot{\gamma} = 0$ otherwise. Note once again that both $\dot{\sigma}$ and $\dot{\gamma}$ are well defined and continuously differentiable in all $\Omega$. 


The convenience of working with the continuous time approximation is clear: we are left with an homogeneous system of two differential equations of order three that is more tractable than the original system of difference equations of order six.

The main result of this paper, formally stated below, shows that if a certain condition on the payoff matrix and/or the learning rule is satisfied, the existence of a significant amount of cooperation is the only stable equilibrium. That is, even if the initial fraction of cooperators is small, the cooperative behavior will grow popular.

**Proposition 3.** If \( f(a, d) > 2f(c, a)f(d, b) \), then the equilibrium \((0, 0)\) is a repeller. On the other hand, if \( f(a, d) < 2f(c, a)f(d, b) \), then the equilibrium \((0, 0)\) is asymptotically stable.

**Proof.** See Lemma 2 in the Appendix.

For a better understating of the conditions in Proposition 3 consider the following example:

**Example.** Assume the payoff matrix of the stage game is given by

<table>
<thead>
<tr>
<th></th>
<th>( C )</th>
<th>( D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C )</td>
<td>( \pi_b - \pi_c, \pi_b - \pi_c )</td>
<td>( -\pi_c, \pi_b )</td>
</tr>
<tr>
<td>( D )</td>
<td>( \pi_b, -\pi_c )</td>
<td>( 0, 0 )</td>
</tr>
</tbody>
</table>

with \( 1 > \pi_b > \pi_c > 0 \). In this case, we can interpret \( \pi_b \) as the benefit a player receives when her partner cooperates and \( \pi_c \) as the cost of cooperating.

**Corollary of Proposition 3.** Assume the stage game is the one given in Table 2 and that players employ the PIR. For any \((\gamma^0, \sigma^0) \in \text{int}(\Omega)\), if \( \pi_b > \pi_c \sqrt{3} \), then a significant amount of cooperation is present in the long run.

The idea behind the survival of cooperation is the following: Imagine a situation where only a small fraction of players cooperate. Some of these players will be matched together, thus, they repeat partner next period. This set of players playing cooperatively and that are matched together obtain the second-maximum payoff, \( a \). Since only very few players cooperate, there is almost no player obtaining the maximum payoff, \( c \). Therefore, under certain conditions, more non-cooperative players imitate cooperative ones than cooperative players imitate non-cooperative ones.
Even if the condition in Proposition 3 is not satisfied, cooperation may still survive if the initial amount of cooperators is high enough. This is proven in our next result.

**Proposition 4.** If \( f(a, d) > f(c, a) \), then the equilibrium \((1, 0) \in \Omega\) is asymptotically stable.

**Proof.** See Lemma 3

To understand the result in Proposition 4, imagine a situation where almost all players cooperate. In this case, if most defectors face other defectors, then cooperative players achieve higher payoff than non-cooperative ones. Thus, under certain conditions, the amount of cooperators increases until all players cooperate. Assume, on the other hand, that most defectors face cooperators. In this situation, defectors achieve higher payoff than cooperators and, thus, the total amount of cooperation decreases. However, the correlation in the matching process favors matches between cooperators and tends to leave defectors matched with other defectors. If conditions in Proposition 4 are satisfied and sufficiently many players cooperate, the payoff from cooperating eventually surpasses that of non-cooperating and the amount of cooperation increases in the population until all players cooperate.

Note that the relationship between the condition in Proposition 3 and the condition in Proposition 4 is ambiguous as \( 2f(d, b) \) can be either greater, smaller or equal to one.

### 4 Robustness Checks and Extensions

#### 4.1 Comparison between Discrete Time and Continuous Time

We have employed the continuous time version of the model to obtain our main result. In what follows, we examine the extent to which this approximation is accurate. Before we present a formal result, the two versions of the model are compared using simulations. Figure 2 shows the simulation of both versions of the model for the same set of parameters as those employed in Figure 1. The value of \( \delta \) is set to 0.001.

As one can see in Figure 2, the continuous time model behaves almost identical to the original discrete time model when close to the boundaries. The fact that the behavior of the two models is almost identical when close to the boundaries is proven below. Since the aim of this paper is to investigate when cooperation survives, the fact that both models converge when \((\gamma, \sigma)\) is close to \((0, 0)\) allows us to conclude that if cooperation survives in the continuous time version of the model then so it does in the original discrete time model and vice versa.
Proposition 5. Define $\Delta x = x^{t+1} - x^t$ with $x \in \{\gamma, \sigma\}$ and let $B_r(\gamma, \sigma)$ be the ball of radius $r > 0$ around the point $(\gamma, \sigma)$. For any $\varepsilon > 0$ and any $(\gamma, \sigma) \in \Omega \cap B_r(0,0)$ we have that if $\gamma^t = \gamma$ and $\sigma^t = \sigma$ then

$$|\dot{\gamma} - \Delta \gamma| \leq o(\varepsilon^2),$$

$$|\dot{\sigma} - \Delta \sigma| \leq o(\varepsilon^2).$$

Furthermore, for any $\kappa > 0$ and any $(\gamma, \sigma) \in \Omega \cap B_r(1,0)$ we have that if $\gamma^t = \gamma$ and $\sigma^t = \sigma$ then

$$|\dot{\gamma} - \Delta \gamma| \leq o(\kappa^2),$$

$$|\dot{\sigma} - \Delta \sigma| \leq o(\kappa^2).$$

Proof. If $(\gamma, \sigma) \in \Omega \cap B_r(0,0)$ then $\gamma, \sigma \leq \varepsilon$. Take equations (5) and (7) and substitute $P_D^\gamma, P_D^\sigma$ and $P_C^{1-\gamma-\sigma}$ from equations (1), (2) and (3). After some algebra, we obtain

$$|\dot{\gamma} - \Delta \gamma| \leq -2\gamma P_D^\gamma + 2\gamma P_D^\sigma + o(\varepsilon^2)$$

$$= o(\varepsilon^2),$$

$$|\dot{\sigma} - \Delta \sigma| \leq -\gamma P_D^\gamma + \gamma P_D^\sigma + o(\varepsilon^2)$$

$$= o(\varepsilon^2).$$
Similarly, when $(\gamma, \sigma) \in \Omega \cap B_\kappa(1, 0)$ we have that $\sigma \leq \kappa$ and $1 - \gamma - \sigma \leq \kappa$. Proceeding as above, we simplify to obtain

$$
|\gamma - \Delta \gamma| \leq -2\gamma P^D_D + 2\gamma P^D_D + o(\kappa^2) \\
= o(\kappa^2),
$$

$$
|\sigma - \Delta \sigma| \leq \gamma P^D_D + \gamma P^F_D + o(\kappa^2) \\
= o(\kappa^2).
$$

\[\Box\]

4.2 Extensions

The long run behavior of the population is determined, to a certain extend, by the initial condition. That is, if no player cooperates initially, then no player ever cooperates. This fact disappears if, for example, mutations or mistakes are introduced in the model. Given that we are dealing with a continuum of population, introducing mistakes is straightforward.

Assume that at any given period with a small probability $\varepsilon > 0$ each player makes a mistake and chooses the action she intended not to. In this case and given that a continuum of population exists, each period exactly a fraction $\varepsilon$ of players make mistakes. More specifically, a fraction $\varepsilon(\gamma + \sigma)$ of players that intended to choose $C$ play $D$, and a fraction $\varepsilon(1 - \gamma - \sigma)$ of players that intended to choose $D$ play $C$.

Results presented are still valid if, in the model with mistakes, an equilibrium is defined as the situation where for any $\varepsilon$ the change in $\gamma$ and $\sigma$ is always smaller or equal than $\varepsilon \gamma$ and $\varepsilon \sigma$ respectively. The convenience of adding mistakes is that unstable equilibria are eliminated. That is, in the model with mistakes, if $f(a, d) > 2f(c, a)f(d, b)$, then cooperation emerges independently of the initial conditions.

5 Conclusions

The present paper investigated cooperation in a setting where players who learn by imitation are matched to play a Prisoner’s Dilemma game. Our contribution to the literature lies in the way matching takes place: players that belong to a pair were both parties cooperated repeat partner while the rest of players are randomly matched into pairs.

In the benchmark case with random matching, we showed that cooperation vanishes for any interior initial condition. When moving to the correlated matching setting, we proved that if mild conditions on the payoff matrix and/or the specific way imitation takes place are
satisfied, then a significant amount of cooperation appears from any initial interior condition. Furthermore, we found that no separating equilibrium exists.

We believe our work is novel in the way the continuous time approximation is employed. This approximation simplifies calculations in models where randomness plays an important role. In our setting, we employed the continuous time approximation to decrease the complexity of the system of difference equations that governs the evolution of the population. We showed that both the discrete time model and the continuous time approximated model converge when the system is close to the boundaries. Thus, if cooperation survives in the simpler continuous time model, then it also survives in the more intricate original model.

REFERENCES


APPENDIX

Random Matching

With random matching, there is no need to distinguish between players who cooperated and were paired with a player who also cooperated, γ, and players who cooperated and faced a player who did not cooperate, σ. Thus, these two sets of players are grouped in the same set \( \omega = \gamma + \sigma \).

Let \( \omega^{t+1} \) the fraction of players who chose C at time \( t \) with \( \omega^0 \in [0,1] \) given. Let \( 1 - \omega^{t+1} \) be the fraction of players who chose D at time \( t \). Furthermore, let \( P_D : A^2 \times R^2 \rightarrow [0,1] \) be the probability with which a cooperative player switches to D and let \( P_C^{1-\omega} : A^2 \times R^2 \rightarrow [0,1] \) be the probability with which a player who chose D switches to C. Assume \( P_D^\omega \) and \( P_C^{1-\omega} \) satisfy Assumptions 1 – 3. The evolution of \( \omega \) is given by

\[
\omega^{t+1} = \omega^t (1 - P_D^\omega) + (1 - \omega^t) P_C^{1-\omega}.
\]  \hspace{1cm} (9)

Proceeding in a similar fashion as in equation (2), we have that \( P_C^{1-\omega} \) is positive only if the player in \( 1 - \omega \) faced another one who played D, and observes an individual choosing C that faced a player who also chose C. In this case, observed payoff equals \( a \) while own payoff equals \( d \). Hence, we can write \( P_C^{1-\omega} \) as follows:

\[
P_C^{1-\omega} = (1 - \omega)\omega^2 f(a, d).
\]  \hspace{1cm} (10)
On the other hand, $P^{\omega}_{D^c}$ is positive only if the player in $\omega$ observed a player in $1 - \omega$. Three different situations can occur now: First, if the player in $\omega$ faced a player in $\omega$ and observed a player who faced another one in $\omega$, then own payoff equals $a$ while observed payoff equals $c$. Second, if the player in $\omega$ faced a player in $1 - \omega$ and observed a player who faced another one in $\omega$, then own payoff equals $b$ while observed payoff equals $c$. Finally, if the player in $\omega$ faced a player in $1 - \omega$ and observed a player who faced another one in $1 - \omega$, then own payoff equals $b$ while observed payoff equals $d$. Therefore, we have that

$$P^{\omega}_{D^c} = (1 - \omega) \left[ \omega^2 f(c,a) + (1 - \omega) \omega f(c,b) + (1 - \omega)^2 f(d,b) \right]. \quad (11)$$

**Lemma 1.** With random matching only $\omega = 1$ and $\omega = 0$ are equilibria. Furthermore, for any $(\gamma^0, \sigma^0) \in \text{int}(\Omega)$

$$\lim_{t \to \infty} \omega' = 0.$$  

*Proof.* We can see from equation (9) that both $\omega = 1$ and $\omega = 0$ are equilibria. The proof is completed by showing that from any point $\omega \in (0,1)$ the system converges to $\omega = 0$.

If $\omega \in (0,1)$, using Assumptions 2 and 3 we obtain the following:

$$\omega^2 f(c,a) + (1 - \omega) \omega f(c,b) + (1 - \omega)^2 f(d,b) \quad > \quad (1 - \omega) \omega f(c,b)$$

$$\geq (1 - \omega) \omega f(c,d)$$

$$\geq (1 - \omega) \omega f(a,d).$$

Thus, we have that

$$\omega f(a,d) \quad < \quad \omega^2 f(a,d) + \omega^2 f(c,a) +$$

$$(1 - \omega) \omega f(c,b) + (1 - \omega)^2 f(d,b).$$

Multiply both sides by $\omega(1 - \omega)$ and use equations (10) and (11) to obtain

$$P^{1-\omega}_{C^+} \quad < \quad \omega (P^\omega_D + P^{1-\omega}_C). \quad (12)$$

From (9) we have that $\Delta \omega = P^{1-\omega}_C - \omega(P^\omega_D + P^{1-\omega}_C)$. Hence, by equation (12), we know that whenever $\omega \in (0,1)$, $\Delta \omega < 0$. Thus, no point $\omega \in (0,1)$ can be an equilibrium and the system cannot converge to $\omega = 1$ from any initial condition $\omega \in (0,1)$.

We still have to show that the system cannot converge to a cycle nor to a point that is not an equilibrium. The fact that the system does not converge to a cycle follows from the observation that for all $\omega \in (0,1)$, $\Delta \omega < 0$. To show that the system cannot converge
to a non-equilibrium point note first that the function $\Delta \omega$ is a polynomial in $\omega$ and, hence, continuous in $\omega$.

Assume the system converges to a non-equilibrium point $\omega \in (0,1)$. Given that for all $\omega \in (0,1)$ we have that $\Delta \omega < 0$, immediately to the right of the point $\omega$ it holds that $\Delta \omega \to 0$, while at $\omega$ it holds that $\Delta \omega < 0$. That is, $\lim_{\delta \to 0^+} \Delta \omega|_{\omega=\delta} = 0$ while $\Delta \omega|_{\omega=0} < 0$. Thus, $\Delta \omega$ is not continuous in $\omega$, a contradiction. $\square$

**Lemma 2.** If $f(a,d) > 2f(c,a)f(d,b)$, then the equilibrium $(0,0) \in \Omega$ is a repeller. On the other hand, if $f(a,d) < 2f(c,a)f(d,b)$, then $(0,0) \in \Omega$ is asymptotically stable.

**Proof.** Define the set $\Sigma = (\gamma, \sigma) \in \Omega \cap B_r(0,0)$. For sufficiently small $\varepsilon > 0$ we can disregard terms of order $o(\varepsilon^2)$ and write the system (7) and (8) when $(\gamma, \sigma) \in \Sigma$ as

$$\dot{\sigma} = -f(d,b)\sigma + f(a,d)\gamma,$$

$$\dot{\gamma} = 0.$$

The approximation above is correct up to a term of order $\varepsilon^2$. Thus, when the process is arbitrarily close to $(0,0)$, the change in $\gamma$ with respect to the change in $\sigma$ is negligible. Furthermore, the system above converges to $\sigma = \gamma \frac{f(a,d)}{f(d,b)}$. Hence, if we start in $\Sigma$ with $\varepsilon$ small, the process converges to a situation where $\sigma = \gamma \frac{f(a,d)}{f(d,b)}$. The system may hit the path $\sigma = \gamma \frac{f(a,d)}{f(d,b)}$ outside the set $\Sigma$. This poses no problem as the further away from $(0,0)$ the system can be in this case is within the set $\Sigma_{\epsilon} \frac{f(a,d)}{f(d,b)}$, which is also arbitrarily close to $(0,0)$ when $\varepsilon$ is small.

After starting in $\Sigma$ and once the system reaches $\sigma = \gamma \frac{f(a,d)}{f(d,b)}$, we can rewrite (8) as

$$\dot{\gamma} = \left(1 - 2\frac{f(d,b)f(c,a)}{f(a,d)}\right)\sigma^2.$$

The equation of the motion of $\sigma$ is irrelevant because in the neighborhood of $(0,0)$ the system moves along the path $\sigma = \gamma \frac{f(a,d)}{f(d,b)}$ as we just proved. To be more precise, the Center Manifold Theorem is being used here (see Sastry (1999) Section 7.8 or Khalil (1995) Section 8.1).

By Bézout’s Theorem, the system (7) and (8) has a finite number of solutions: a system of two equations, two unknowns of degree three has at most 6 solutions (see Kirwan (1992)). Thus, we can fix $\varepsilon > 0$ such that no equilibrium points exists in $\Sigma_e \setminus (0,0)$. Assume now that $f(a,d) > 2f(c,a)f(d,b)$ so $\dot{\gamma} > 0$ in $\Sigma$. Local stability implies that for any $\varepsilon > 0$ we
can find a $\kappa < \varepsilon$ such that if the system starts in $\Sigma_\kappa$, then it never leaves $\Sigma_\varepsilon$. Assume this is the case.

For any $\kappa < \varepsilon$, if $f(a,d) > 2f(c,a)f(d,b)$ then $\dot{\gamma} > 0$. Thus, since $\dot{\gamma} > 0$ and $\sigma = \gamma \frac{f(a,d)}{f(a,d) - f(d,b)}$, if the system starts in the boundary of $\Sigma_\kappa$, then it will leave that set. Assume that the system, after leaving $\Sigma_\kappa$, does not hit the boundary of the other bigger set $\Sigma_\varepsilon$. Since for any point in $\Sigma_\varepsilon$ we have that $\dot{\gamma} > 0$, by continuity of (7) and (8) if the process does not hit the boundary of $\Sigma_\varepsilon$ then we must have that there exists a point $(\gamma, \sigma) \in \Sigma_\varepsilon \setminus (0, 0)$ such that $\dot{\gamma} = 0$ and, thus, $\sigma = 0$. That is, there must exists at least one equilibrium point in $\Sigma_\varepsilon \setminus (0, 0)$, which is a contradiction.

Thus, if the process starts in $\Sigma_\kappa$, then it must hit the boundary of $\Sigma_\varepsilon$. We know that for any point in $\Sigma_\varepsilon$, if $f(a,d) > 2f(c,a)f(d,b)$ then $\dot{\gamma} > 0$ and $\sigma = \gamma \frac{f(a,d)}{f(a,d) - f(d,b)}$. Thus, starting in boundary of $\Sigma_\kappa$ the process leaves $\Sigma_\varepsilon$, which is the condition for the point $(0, 0) \in \Omega$ to be a repeller.

Assume now that $f(a,d) < 2f(c,a)f(d,b)$. By continuity, $\dot{\gamma} < 0$, $\sigma = \gamma \frac{f(a,d)}{f(a,d) - f(d,b)}$ and the fact that no equilibrium point exists in $\Sigma_\varepsilon \setminus (0, 0)$, if the system starts in $\Sigma_\varepsilon \setminus \Sigma_\kappa$ then it eventually enters the set $\Sigma_\kappa$ for any $\kappa < \rho$ with $\rho < \frac{f(d,b)}{f(a,d)}\varepsilon$. This is the condition for asymptotic stability.

\begin{lemma}
If $f(a,d) > f(c,a)$, then the equilibrium $(1,0) \in \Omega$ is asymptotically stable.
\end{lemma}

For proving Lemma 3 we employ Lyapunov’s method:

\begin{proposition}[Theorem 4.7 in Khalil (1995)]
Let $x = 0$ be an equilibrium point for a system described by

$$x = f(x)$$

where $f : U \to \mathbb{R}^n$ is locally Lipschitz and $U \subset \mathbb{R}^n$ a domain that contains the origin. Let $V : U \to \mathbb{R}$ be a continuously differentiable, positive definite function in $U$ with $V(0) = 0$.

- If $\dot{V}$ is negative semi-definite, then $x = 0$ is a stable equilibrium point.
- If $\dot{V}$ is negative definite, then $x = 0$ is an asymptotically stable equilibrium point.
\end{proposition}

\begin{proof}
Consider the set $\Omega$ and take the system given by (7) and (8). As argued in the main text, this system is continuously differentiable and, thus, Lipschitz continuous in $\Omega$.

Lyapunov’s method easily applies to the equilibrium $(1,0)$ by requiring $V$ to be such that $V(1,0) = 0$. Take $V(\gamma, \sigma) = 1 - \gamma - \sigma$.
It is easy to see that $V(1, 0) = 0$ and $V(\gamma, \sigma) > 0$ for points in $\Omega$ around $(1, 0)$. We have to show that $\dot{V}(\gamma, \sigma) < 0$ for $(\gamma, \sigma) \in \Omega$. From (7) and (8), the function $\dot{V}$ is given by

$$\dot{V}(\gamma, \sigma) = -\frac{1 - \gamma - \sigma}{1 - \gamma} \left[-\gamma \sigma f(c, a) - \sigma^2 f(c, b) - (1 - \gamma - \sigma)\sigma f(d, b) + (1 - \gamma - \sigma)\gamma f(a, d)\right].$$

When the system is close to $(1, 0)$, we can rewrite the expression above as

$$\dot{V}(\gamma, \sigma) = \frac{1 - \gamma - \sigma}{1 - \gamma} \left[\sigma f(c, a) - (1 - \gamma - \sigma)f(a, d)\right].$$

Thus, given that $\sigma \leq 1 - \gamma - \sigma$, if $f(a, d) > f(c, a)$ then $\dot{V}(\gamma, \sigma) < 0$ around $(1, 0)$ and, hence, by Proposition 6 the point $(1, 0)$ is asymptotically stable. \qed