Cooperation as self-interested reciprocity in the Centipede

Francesco Farina and Patrizia Sbriglia

university of siena -university of Naples II

February 2007

COOPERATION AS SELF-INTERESTED RECIPROCITY IN THE CENTIPEDE[§]

Francesco Farina
University of Siena
farina@unisi.it

Patrizia Sbriglia[ξ]
University of Naples II
patrizia.sbriglia@unina2.it

February 2007

Preliminary version

Abstract

Cooperation is a pervasive social phenomenon but more often than not economic theories have little to say about its causes and consequences. In this paper, we explore the hypothesis that cooperative behaviour might be motivated by purely selfish interest when the "social" payoff in a game is increasing. We report the results of a series of experiments on the centipede game. The experiments are organized in two subsequent steps. Subjects first participate in a 2-period *trust game*, randomly matched with unknown partners. We apply the strategy method in order to elicit their social preferences. On the basis of their pre-game behaviour, individuals are divided into three main social groups: *selfish individuals, pure altruists* and *reciprocators*. At the second step of the experiment, subjects play a repeated 6-move centipede game with increasing final payoff. Each subject plays twice in a low stake and in a high centipede game, and he/she is informed about his/her co-player social preferences. We identify the origin of cooperation within homogeneous and heterogeneous social groups.

Keywords: *social preferences, altruisms, experiments.*

J.E.L. classification: C91,C92.

[§] We thank R.Ricciuti, for his valuable suggestions and contributions to the Instructions of the experimental sessions, and F. LoMagistro for his excellent technical assistance. We also thank E. Ciriolo, A. Innocenti e N. O'Higgins for helpful comments on a previous version of the paper. Financial support from the research programme of the University of Siena is acknowledged. The usual disclaimers apply.

[ξ] Corresponding author: Patrizia Sbriglia, Department of Law and Economics, University of Naples II, Via Gran Priorato di Malta, Capua, Caserta, 83043, Italy, tel: 00390823274355; fax: 00390823628694; e-mail addresses; patrizia.sbriglia@unina2.it; sbriglia@unina.it.

COOPERATION AS SELF-INTERESTED RECIPROCITY IN THE CENTIPEDE

1. Introduction

Theoretical paradoxes challenging the axiomatic foundations of *Expected Utility Theory* (EUT) show systematic deviations of agents from selfish behaviour. The tendency to consider the self-regarding individual as the "representative" agent in economic modelling has been scrutinised and found wanting. In a seminal paper, a reciprocator was defined as an individual who responds fairly to fair behaviour and with hostility to hostile behaviour (Rabin, 1993). This preference for conditional cooperation was successfully verified in experiments. Agents seemed to be willing to cooperate, provided that other individuals involved in social interactions were equally committed.

The most substantiated approach in the experimental literature interprets reciprocating behaviour as the desire to preserve the fairness conditions of social interactions (Fehr and Schmidt, 1999, 2002, 2006). In Ultimatum and Dictator games, a player responds fairly to a fair offer, and punishes whenever he feels that an other-regarding behaviour has been waived in favour of a selfish one (Fischbacher, Gachter and Fehr, 2001; Falck, Fehr, and Fischbacher, 2003). Similarly, the presence of a "critical" number of individuals motivated by an attitude of reciprocation, interacting with a multitude of "straightforward maximizers", is considered crucial in making public goods provision possible (Bolton and Ockenfelds, 2001; Fehr and Fischbacher, 2002).

A feature of this approach is the view that the incentive to "reward or punishment" mainly reflects a taste for "strong reciprocity" in one-shot interactions, as opposite to the "weak reciprocity" which characterizes the self-interested behaviour of individuals involved in long-term relationships which are typical of repeated games (Fehr and Schmidt, 2006). Apart from the "pure altruist" - the individual who is disposed to cooperate unconditionally, even at the cost of a personal loss – this approach devises a *continuum* of conditional behavioural dispositions. The possible types range from the inequity aversion of the "conditional altruist" – whose utility increases as a result of the distribution becoming more equitable towards another individual – to the commitment to "conditional cooperation" – the will to establish and/or preserve a social norm of fairness, prevailing on the search for the maximum pay-off in strategic interactions.

Therefore, the behaviour of the conditional cooperator should not be associated with the appearance of cooperation among selfish players aimed at Pareto optimization: "It is important to emphasize that reciprocity is not driven by the expectation of future material benefit. It is, therefore, fundamentally different from 'cooperative' or 'retaliation' behaviour in repeated interactions (…); in the case of reciprocity, the actor is responding to friendly or hostile actions *even if no material gains can be expected* (italics added)" (Fehr and Fischbaker, 2002). This account of conditional cooperation is not far different from the definition of "altruism", as proposed by this philosopher Thomas Nagel: "By altruism I mean not abject self-sacrifice, but merely a willingness to act in the consideration of the interest of other persons, *without the need of ulterior motives* (italics added)" (1970, p.79). From this vantage point, the conditionality of cooperation as compliance with a social norm of fairness is sharply opposed to the selfish behaviour of the individual who is disposed to cooperate on the basis of pay-off maximization.

In this paper, we conduct experiments by setting up an analytical framework which to some extent moves away from this view. We argue that a too thick wall has been built dividing "conditional cooperation" from "self-interest". The Folk Theorem relates cooperation in infinitely PD-like repeated games to the set of sub-game perfect equilibria containing all strategies with an average pay-off not inferior to the one that could be gained in case the opponent struggles to spoil you. This game-theoretic assessment of the rationality of instrumental reciprocity conceals the inescapable ambiguity which is embedded in the conceptualization of cooperation among selfish players. Even in one-shot games, an act against selfish behaviour and in favour of fair behaviour may be motivated by the desire to establish and/or preserve a social norm of fairness, but in combination with the desire to mutually insure higher pay-offs in possible future social interactions: "Individuals forgo their short-term selfish gains because being nice (or, more precisely and more generally, playing their equilibrium strategy) will lead to nice treatment in the future. Punishing nasty behaviour serves to discourage nasty behavior, but punishment only occurs because players fear that a failure to punish will lower their future payoffs" (Sobel, 2004, p.30). Similarly, the possible interplay between altruism and self-interest has been modelled as the individual behaviour motivated both by self-esteem *and* by awareness that building up a social reputation leads to higher returns (Benabou and Tirole, 2004).

This more balanced approach has also been advocated by Jon Elster: "I believe that both norms and self-interest enter into the proximate explanations of action. To some extent, the selection of the norm to which one subscribes can also be explained by self-interest. Even if the belief in the norm is sincere, the choice of one norm among the many that could be relevant may be an unconscious act dictated by self-interest Or one might follow the norm out of fear of the sanctions that would be triggered by violation. But I do not believe that self-interest provides the full explanation for adherence to norms" (Elster, 1998).

The focus of our paper is on the interdependence between the homogeneous/heterogeneous types of the agents and the specific strategic situation in which they interact. We conduct experiments in which two players, randomly matched, first play a one-shot Trust game, and then are involved in the multi-stage joint-venture formalised by the Centipede (Rosenthal, 1981), with (or without) the information about their opponent's behaviour in the pre-game.

2. The choice of the Centipede

The motivation of the choice of the Centipede for our experimental enquiry about the origins of reciprocal behaviour is that this game considers the incentivse for reciprocity in the context of expectations of increased pay-offs, that is the two players trusting each other to benefit from the mutual advantage of continuing the game, possibly till the final stage. We underline the central role played by the information putting each player in the condition to rationalise his opponent's expected behaviour. Before presenting the set up of the experiments, it is worth clarifying that this stress on information finds its theoretical underpinnings in the defence put forward by Binmore to justify the violation of backward induction in the Centipede.

Aumann (1998) offered the formal proof of backward induction dictating the first mover to play *down* at the first stage of this game. Yet, experimental evidence shows that a not negligible percentage of players move *across* at the first node (McKelvey and Palfrey, 1992). The second player's decision to play *across* as well is even harder to rationalize. Since "there cannot be any mistakes if the players are absolutely rational" (Selten, 1975, p.25) and do not make mistakes, an explanation is lacking for this behaviour of reciprocation in a certain number of stages of the experiments conducted on the Centipede.

The inconsistency between common knowledge of rationality (CKR) dictating this game to be over at the first node, and evidence showing the violation of the Nash equilibrium solution, provoked a dispute between Binmore (1988, 1996) and Aumann (1995, 1996). By taking issue with the Aumann (1995) appraisal of rationality in the Centipede Game, Binmore (1996) has advocated the need to consider the players' Bayesian updating at each node they reach[1]. He endorsed the view that the infringement of backward induction has to be taken as a signal launched to the other player, in order to both keep on playing *across* for the whole duration of the game but the last node. In his rejoinder, Aumann dismissed as

[1] Binmore argues that, whether or not CKR is assumed, out-of-equilibrium moves are a substantial part of our understanding of games such as the Centipede. "(P)layer I can quantify his ignorance about what would happen if he were to play *across* (and) assign a probability p to the event that the result of his playing across would be a payoff of at least 4, rather than the payoff of 2 he gets by playing *down*. The latter eventuality would result, for example, if Player II were to deduce from Player I's choosing *across* at the first node that Player I would also choose *across* if the third node were reached" (Binmore, 1996, p.136). In his rejoinder, Aumann sticks to his point whereby, under CKR "in deciding whether a player i is rational when choosing down in the centipede game, we do explicitly take into account what i knows or thinks about what the other player would have done if i had gone across" (pp.140-1).

"absurd", in the absence of CKR, to argue that "the beliefs of a rational player might motivate him to play *across*" at a node of the Centipede (Aumann, 1996, pp.141-2). His resolute denial of any legitimacy of a recourse to beliefs and probabilities according to Bayesian rationality reflects the impossibility of out-of-equilibrium moves in the context of orthodox rational choice theory. Since the player's strategy consists of the maximisation of the expected playoff at *every* information set, equilibrium strategies should be self-enforcing also at out-of-equilibrium nodes.

To impose that what would happen if *across* were played by the first player cannot be evaluated, amounts to suggesting that beliefs do not change even though out-of-equilibrium deviations occur in the pursuit of higher payoffs. Hence, the hypothesis of CKR cannot be valid not only throughout the game but even as a "prior hypothesis", that is before the beginning of the game. Once the CKR hypothesis has been cancelled out by the first mover playing *across* in a four-nodes version of the Rosenthal's Centipede, Reny (1992) has demonstrated that the common beliefs of rationality (CBR) applies at every node if the sub-game perfect equilibrium play reaches *all* nodes, out-of-equilibrium nodes included. This theoretical appraisal confirms the intuition of Rosenthal (1981), who argued that the Centipede players assign at each node a positive probability to their opponent's decision to play *across*. In the same vein, the McKelvey and Palfrey (1992) experiments shows that perfect information on the length of the Centipede is important as well, as players reach more distant nodes the longer is the structure of the experimental game.

Therefore, the structure of the Centipede is particularly suited to combine two possible mechanisms for conditional cooperation which may be at work: the attitude to reciprocity as an effect of the different degrees of trustworthiness arising in different mixtures of types, and the strength of the incentive to maximization depending on different levels of pay-offs. In conducting our experiments, we aim to show the importance of knowledge of the opponent's set of opportunities[2]. The possibility to set up informed expectations about the other player's behaviour at each node of a multistage social interaction is taken as the pre-condition of trustworthiness between players, which is instrumental to the insurgence of common social preferences.

3. The Experiment

Our experimental design consists of a Trust game followed by a Centipede game. We exploit the connection between the respective strategic environments, as the latter is a sort of a multi-stage Trust game and the former is a one-stage Centipede. After agents have played as Senders and Respondents in the Trust Game, their behaviour is observed in a six-stage version of the Centipede, with and without the information elicited in the pre-game about the opponent's reciprocating strategies. The rationale for this blend is that the Centipede mimicks a one-shot market game in which the agents' behaviour is characterised by complete information on strategies and payoffs, but uncertainty about the types of the players. In contrast to public goods games, the choice of a reciprocating behaviour depends on the expectation about the other agent's behaviour. The information gathered in the experiment conducted with the Trust game is instrumental in letting the Centipede players set up a probability for opponents' behaviour in each stage of the game, and thus their strategy in the experiment. The pre-game represents the social environment which allows players in both position A and position B of the Centipede to rationalise the out-of-equilibrium nodes of the extensive form game by relying on what the other individual declares to be willing to send back when playing as a Responder.

We analyse whether in this extensive-form game reciprocal behaviour becomes more likely – and possibly proceeds node after node - when matching us between two players whose types are homogeneous/heterogeneous and/or with a "high-stake"/"low-stake" payoff

[2] The role of the agents' types in setting up a pattern of reciprocity is stressed by Levine (1998, p.620).

structure of the game. First, we compare two sets of sessions - with players endowed first with incomplete, and then complete, information about the social preferences and increasing pay-offs – in order to find out which framing is more effective in establishing reciprocity between self-regarding individuals. Second, when the Centipede is played with complete information, the players are divided into categories each one corresponding to a type, in order to check whether the propensity to cooperate is higher when a homogeneous or a heterogeneous matching between players occurs.

The rationale we would like to put forward is that Centipede players reciprocate not because of a social norm or because of mutual advantage, but due to a blend of both motivations. On the one side, matching is decisive, as each player becomes knowledgeable about similar reciprocating attitudes of the other player and is able to learn cooperative behaviour by joining the other individual's decision to cooperate. On the other side, the players are able to exploit the peculiar feature of this game in extensive form - to assign to each player, in alternation node-by-node, a lower and an increased payoff – by letting the crop grow in order to maximize their expected utility at each stage, till the last node where also a conditional cooperator last mover is expected to free-ride. Due to the mutual advantage permitted by the favourable environment reflected by the increasing collective payoffs, reciprocal compliance allows players to cooperate in improving their respective individual welfare. Instead of the Pareto-inefficient result in which the two players receive the lowest collective payoff, the Pareto-optimal collective payoff might be obtained.

The paper is organised as follows. In Section 4, we present the experimental design and the financial incentives; our hypothesis testing is stated at the end of the Section. Section 5 contains the analysis of the results of our experiments. This section is divided into three parts. Firstly, we try to assess the importance of information on the subjects' reciprocating behaviour by comparing the level of cooperation in the four sessions of the Centipede games in which such information was available, with the level of cooperation in our baseline setting in which players played two rounds of the Centipede with equal payoff structures, but with no information on opponents' attitudes. We secondly analyse the pre-play experimental evidence on the Trust games and we classify the social behaviour of the participants. Finally, Section 6 reports the study of the cooperation rates in homogeneous and heterogeneous groups and analyses the individual incentive to choose the "Continue" strategy, in homogeneous and in heterogeneous pairs.

Section 6 concludes and offers few suggestions on the directions along which our research may be expanded.

4. Experimental Design, Financial Incentives and Experimental Predictions.

The experiments were conducted in Siena (October-December 2006) and subjects were recruited among undergraduate and graduate students of Law, Business and Economics. Participants were paid according to their cumulative profits and earnings varied between 12 and 16 Euro per subject. In total, 100 students participated in our experiments, divided into five groups of 20 subjects per sessions. We ran our experiments with two different designs. In the first design, composed of one session (denoted Session 0 in the Tables), subjects were coupled randomly and they were asked to play two rounds of the following six-nodes centipede games :

GAME 1 (Low Payoff):

A → **C** → B → **C** → A → **C** → B → **C** →A → **C** → B → (160) (130)

E ↓	**E** ↓	**E** ↓	**E** ↓	**E** ↓	**E**↓
(40)	(30)	(80)	(70)	(120)	(110)
(10)	(60)	(50)	(100)	(90)	(140)

GAME 2 (High Payoff):

A → **C**→ B → **C** → A → **C** → B → **C** → A → **C** → B → (260) (210)

E ↓	**E** ↓	**E** ↓	**E** ↓	**E** ↓	**E** ↓
(140)	(110)	(180)	(150)	(220)	(190)
(90)	(160)	(130)	(200)	(170)	(240)

where A moves first and the choice between strategies E and C implies that the game ends (E), or that the game continues (C). Payoffs were expressed in experimental tokens and the exchange rate was fixed to 1 Euro cent per token and reported in the Instructions.[3]

The second experimental design consisted of 4 sessions (denoted as Session 1 to Session 4 in the Tables) and were organized in two subsequent steps. Subjects first participate in a 2-period *trust game* (TG, hereafter), randomly matched with unknown partners. They were then asked to participate in the two rounds of the Centipede games as in Session 0.

However, when acting as Respondents' in the TG, we applied the strategy method in order to elicit their reciprocating preferences. As the first step in the game, in fact, players were asked to report – for each number of tokens received by the Sender – how many tokens they would return. The initial number of tokens allocated to the Sender was set to 10 units and the multiplying factor was fixed to 3 and reported in the Instructions.

The second step consisted in reporting the number of tokens they would send, when playing in the Sender role, then the computer selected random pairs of Respondents and Senders and players collected the payoffs earned when playing in both roles.[4]

The reason why we adopted the random matching and anonymous partnership protocols is that, in our opinion, they would enhance the incentive of a true revelation of the individual preferences.

An important point of the experimental design of the pre-play game is the choice of strategy method (Selten, 1967) as the procedure to elicit the individuals' other-regarding preferences.

Several experimental methodologies have been used to measure individualistic, reciprocating or cooperative and altruistic behaviours. We recall here the use of questionnaires, pre-play one shot or repeated games and finally some variations of the

---

[3] A copy of the Instructions is available on request.

[4] The structure of the Trust Game we adopted in our experiments is known as the Generalised Trust Game (hereafter, GTG), in the sense that Respondents sent tokens back not to the Sender who sent to them, but to some other anonymous Sender present in the same session (see Barr *et al.,* 2005).

strategy method (see Burlando and Guala, 2005, for extensive references). Each of these methodologies has been criticised on several grounds. In the case of the strategy method, possible disadvantages are related to the weakening of incentives , since each state of the world occurs with less than unitary probability and problems of cognition and understanding may arise, as the number of observations on the players' (in our case, the Respondents) behaviour increases (in our case, Respondents were asked to indicate 10 values of the number of tokens they would return to the Sender). Finally, according to some authors (Guth *et al.* 2001), the strategy method may have an impact on individuals' social preferences, thus weakening the validity of its application as a mean to classify reciprocating behaviours. In our opinion, however, similar remarks may be made about the methodologies of the one-shot and the repeated pre-play games, whilst, in the case of the questionnaires, the reliability of the answers may be questioned.

Furthermore, the strategy method has the important advantage of providing each player with a wide representation of the other player's choices, motivations and social tendencies. In this respect, since the aim of our work is to study the level cooperation in Centipede games, in a context where players have full information on their co-players social types, we believe that this specific methodology is the most apt to answer our research questions.[5]

Once the TG was played and payoffs were distributed, all subject would receive the second set of Instructions relative to the Centipede games. The experiments lasted 12 periods. At the beginning of each period, the computer matched the subject with an unknown partner. Then, they were randomly given the role of A or B; each player playing at some stage however in both roles. Before the actual game started, each person was shown the table previously filled

[5] In Gachter and Thoni, 2004, players were informed on the social types they were playing with in repeated public good games. The information on the reciprocating strategies was elicited by a pre-play participation of the subjects in a one shot public good game. Subjects were then ranked according to their performance in the pre-play game, and such information was disclosed to the other players when participating in the experiments. In our opinion, this methodology provides only very limited information on social preferences. Moreover, it does not address the problems of understanding and possible impact of the method itself on the individuals' preference functions.

out by his/her co-player. The computer allowed a minute to look at the table, and the game then started.[6]

As reported in the Instruction sheet, the only information available to players was individuals' reciprocating strategies as reported in the Respondents' Table, and not the Senders' choices, since we focus primarily on the effect of trustworthiness and reciprocation on the cooperative levels in the centipede games.

5. Results.

In order to address the issues raised in the introduction, we analyse the experimental evidence in three subsequent steps.

Firstly, we compare the level of cooperation in Session 0 and Sessions 1 to 4. Our aim is to assess the role played by the information on social preferences on the cooperation levels.

Secondly, we analyse the experimental evidence gathered in the TG games, in order to classify "social types", so that their behaviour can be monitored when examining the data in Session 1-4. We then study the behaviour of groups (pairs) and individuals in Session 1-4, and see if there are significant differences between cooperation levels in homogeneous and heterogeneous groups.

Finally, we look at individual behaviour and try to assess – for each "type" - the incentives to cooperate in the Centipede games.

[6] It may be argued that the Respondents' Table, composed by ten observations, was quite a complex piece of information to understand in a limited amount of time. There are however other experimental designs, which provide complex information on decision variables, such oligopoly experiments played as normal form games (see Holt, 1995).

5.1 Cooperation in Session 0 and Sessions 1-4

Our first objective is to assess the importance of the information on social preferences on the choices of the "E" and "C" strategies in the Centipede games. To this end, we will compare the evidence gathered in Session 0 and Sessions 1-4.

Tables 1 and 2 report the frequencies of the E strategy at each node and for both players A and B, for the low and high payoff games.[7]

TABLE 1: Frequencies of the Exit Strategy in the Low Payoff Centipede Games

Player A Nodes	**Session 0**	**Session 1**	**Session 2**	**Session 3**	**Session 4**
1	0.08	0.14	0.09	0.10	0.15
3	0.13	0.14	0.18	--	--
5	0.23	0.26	0.63	0.60	0.69
7	0.56	0.46	0.09	0.30	0.16
Player B Nodes					
2	0.08	0.08	0.10	0.15	0.23
4	0.23	0.18	--	0.31	0.22
6	0.23	0.28	0.80	0.31	0.33
8	0.46	0.46	0.10	0.23	0.22

[7] For expositional purposes, we report the final node as node 7, for player A, and node 8 for player B.

TABLE 2: Choices of the Exit Strategy in the High Payoff Centipede Games

Player A	Session 0	Session 1	Session 2	Session 3	Session 4
1	0.25	0.50	0.28	0.27	0.23
3	0.33	0.33	0.27	0.46	0.31
5	0.25	0.17	0.45	0.27	0.38
7	0.17	--	--	--	0.08
Player B					
2	0.30	0.43	0.28	0.11	0.13
4	0.30	0.43	0.63	0.67	0.63
6	0.30	0.14	0.09	0.22	0.12
8	0.20	--	--	--	0.12

The Tables allow us to address two different issues, namely the effect of information on the incidence of the Nash equilibrium in both types of games, and the overall impact that information have on behaviour.

Let us begin with the relative frequency of choices which correspond to the Nash equilibrium (i.e., the choice of E at the first node for A and the choice of E at the second node for B), and , conversely, the relative frequency of choices which correspond to the fully cooperative outcome (i.e., the choice of C at the fifth node for A and the choice of C at the sixth node for B).

Tables 1 and 2 show that when there is information on the social preferences of the co-players, there is an overall decrease in cooperation.

Specifically, both for the low and high payoff games, the frequency of "exit" at the final node is lower in Session 1-4 than in Session 0. In the low stake games, the proportion of A players choosing to exit the game at the seventh step drops from 56 per cent in Session 0, to a minimum of 9 per cent (Session 2). By the same token, the proportion of B player exiting the game at the final node varies between 46 per cent in Session 0 and the minimum of 10 per cent of Session 2.

The negative effect of information on the proportion of players choosing the full cooperative outcome is even greater in the high stake than in the low stake games.

In three cases out of four (Sessions 1, 2, 3) the proportion of A (B) players exiting the game at the final node drops from 17 per cent (20 per cent) to zero; whilst in the Session 4 the proportion of A (B) exiting the game at the final node is equal to 8 per cent (12 per cent).

On the contrary, in the presence of information on the reciprocating attitudes of the co-players, the incidence of the Nash equilibrium choices tend to increase both in the high and low stake games. In fact, when there is full information, with only one exception (Session 4, high stake game), the proportion of A's playing "E" at node 1 is higher in Session 1-4 than in Session 0; moreover, in one case (Session 1, high payoff centipede games) such proportion is twice as much as in the baseline design.[8]

It must noticed that the proportions of the Nash equilibrium choices in Session 0 are in line with the previous studies on the six-moves Centipede game (see Camerer, 2003; Nagel and Tang, 1998), and, therefore, it is reasonable to argue that there is a clear effect of the information on the individuals' reciprocating strategies on the Nash Equilibrium choices in the low stake and in the high stake settings.

In addition to the exam of how information affected the convergence to the equilibrium points, we are interested in assessing whether it had a more general influence on behaviour

[8] The proportion of B players exiting at node 2 in Session 1-4 is higher than the same proportion in Session 0 in four cases out of eight.

and therefore we are interested in analysing the individuals' choices along the entire game tree.

Confining the attention to the A players, the Mann-Whitney U-tests on the comparison between the frequency distributions of "E" in Session 0 and in Session 1-4 (aggregated in one session) reject the null hypothesis of no difference in choices both in the low payoff games (p=0.10) and in the high payoff games (p=0.25). This implies that individuals reacts to the information not only exiting the game earlier than in the alternative experimental settings, but they change their course of actions, probably in response to the specific social type they are facing.

5.2 Classification of "Social Types" in Sessions 1-4

The comparison between Session 0 and Sessions 1-4 has thus far provided an aggregate picture of the effects of information on reciprocating strategies on cooperation levels in the Centipedes. A further step is to examine the levels of cooperation for pairs of players, with similar or different reciprocating profiles.

For this reason, in what follows, we concentrate our attention on Sessions 1-4, starting with the analysis of the pre-game findings.

Figures 1-4 report the Responders' average behaviour in each of our four GTG games, that is, for each possible value of the tokens sent by the Proposers (3, 6, 9, etc.), each figure reports the total average number of tokens returned by Responders in that specific session. Figures 1-4 show that the Responders' strategy profiles are close to what is defined as "implicit" and "conditional" reciprocation (Sobel, 2004, Fischbaker *et al.*, 2001), i.e., the willingness to be generous to generous players and mean to mean players[9]. Overall, however, reciprocation – albeit unfair reciprocation - is a widespread code of behaviour.

9

- INSERT FIG. 1-4 ABOUT HERE –

Let us begin considering the individual average behaviour in the pre-game GTG sessions. More than 90 per cent of the individuals' strategies responses can be classified as (unfairly or fairly) reciprocating behaviour and only 5 individuals out of 80 declared to have purely selfish social preferences. If pure selfishness is a rare social characteristic, pure altruism is an even rarer one. No participants can be classified as pure altruists, but 32 per cent of the reciprocating behaviour can be classified as 'more than fairly' reciprocating behaviour and in two out of 75 cases, social preferences are closer to purely altruistic behaviour than to fairly reciprocating behaviour. The incidence of "flat" strategy types (i.e. reciprocating strategy profiles with zero variance: players returning fixed amounts of money regardless of the Sender's behaviour) is also limited (two individuals in our sample), though many unfairly reciprocating profiles possess a low variance of choices, showing that unfairness is often independent of the Senders' actions. We have three pure "fair types", but "almost" fair behaviour can be found in 8 cases out of 80.

In a more general way, we need to consider whether the individual behaviour can be classified as *selfish, altruistic or fair* social behaviour. We implement such classification by examining a statistical measure of the individuals' distributions as reported in their Respondents' table, represented by the individual's mean deviation from the theoretical fair reciprocator type, MD:

$MD = \sum_{i=1}^{20} \left(\frac{x_i - \bar{x}}{(n-2)\bar{x}} \right)$; where $\bar{x}$ is the theoretical fairly reciprocating distribution,[10] while x_i is the individual's observed distribution as reported in his/her strategies' profile.[11]

[10] Such distribution corresponds to the equal split of the total number of tokens between Responders and Senders, in the eight cases (out of ten) in which such equal split can be autonomously implemented by the Responder.

Our sample data can be therefore divided in three groups of different "social types":

1. Selfish players: MD < 0; we denote these players as $\boldsymbol{\alpha}$.
2. Fair players MD = 0; we denote these players as $\boldsymbol{\beta}$.
3. Altruistic players MD > 0; we denote these players as $\boldsymbol{\gamma}$.[12]

5.3 <u>Groups and individuals: the study of cooperation in the low and high payoff Centipedes</u>

In Tables 3-6, we report, for each session and each pair of players, their type, the type of their co-player, and the "exit node", both in the low stake and in the high stake games.

The Tables provide a first insight on the level of cooperation per groups and individuals in our experiments. However, we will proceed to analyse the findings in the low and in the high stake games, subsequently. In total, in the low stake Centipedes, the sample data were split into 38 groups with homogeneous preferences and 42 groups with heterogeneous preferences, whilst, in the high stake framework, the sample was more unequally split into 32 pairs with homogenous preferences and 48 with heterogeneous ones.[13]

---

[11] In a similar fashion, Barr *et al.*2005 apply the strategy method in the GTG to elicit the individuals' social preferences. In their paper, the authors focus on several hypotheses on how Respondents' behaviour may be classified in the GTG. The main hypothesis tested is that the latent reciprocating strategies adopted by the Respondents are rays, i.e., "they are linear in the amount sent by Proposers and pass through the (nothing sent, nothing returned) origin" (p. 16).
According to this hypothesis, social behaviour may classified as selfish, reciprocating and altruistic, according to the estimated slope of the individuals' reciprocating strategies' profiles. As an example, selfish strategies are linear strategies with zero slope. Reciprocating strategies are monotonically increasing strategies with the value of the slope, β, varying in the interval: $0 < \beta \leq 3$, where 3 is the value of the multiplying factor, α. We do not adopt this methodology and prefer to follow simple criteria as the convergence of the individual preference profile to a theoretical type.

[12] Overall, 48 individuals in our sample belong to the first category, 8 to the second and 24 to the third one.

[13] There are two aspects that need to be underlined in order to assess the relevance of our results. First, both samples do not contain homogeneous pairs of the β type (the reciprocator players). Second, because of the larger number of selfish players in our experiments, the sample data of the homogeneous pairs is constituted by a higher proportion of selfish players, relative to altruists. Since both aspects may partly explain the high level of cooperation we observe in the heterogeneous groups, we will analyse the evidence of groups and individuals conjointly, in order to evaluate the relative incentive to cooperate of each social type, both when they face a similar type of player and when they face a different type.

Figures 5 and 6 report the proportions of the exit choices at each node in the Centipedes, both for the heterogeneous and the homogeneous groups, in the low and high stake frameworks. In both cases, we can observe that heterogeneous groups tend to be more cooperative than homogeneous ones. There are however differences between the two games that need to be underlined. Specifically, in the low stake game, cooperation is very high for both groups, and there is not a large difference between the choice of “exit at the first (and second) stage” between homogeneous and heterogeneous pairs. In fact, the cumulative frequency of the “E” strategy in steps 5-7 is equal to 58 per cent in the homogenous group and to 71 per cent in the heterogeneous ones. By the same token, the frequency of the choice of “E” at step 1-2 is equal to 20 per cent for the former group and 10 per cent for the latter one.

On the contrary, there is a sharp difference in behaviour in the high stake game. The cumulative frequency of the “E” strategy in steps 5-7 is equal to 18 percent and 28 per cent, respectively for homogeneous and heterogeneous pairs, while the “exit” at the first and second nodes correspond to 40 per in the first group and 20 per cent in second.

The difference in behaviour across groups is confirmed by the χ^2 statistical test aimed to compare the frequency distributions of the homogeneous and heterogeneous pairs in two settings. We tested the hypothesis that the behaviour of the subjects was consistently different in homogeneous and heterogeneous couples (i.e., the two distributions are independent). Both in the low and high stake games , the χ^2 test accepted the null hypothesis (5 per cent significance level).

The overall higher level of cooperation we register in the heterogeneous cases posits an important question in relation to the individuals’ incentives to cooperate: do altruists (selfish players) change their behaviour when facing selfish players (altruists)?

Table 7 tries to provide an answer to the previous question. In the Table we report the frequency of the “C” strategy per couple, identifying the different social preferences of each individual. Comparing the values of the frequencies among homogeneous and heterogeneous pairs, we can notice that the largest difference in cooperation rates is registered when two selfish players interact, in the low and in the high stake games.

High levels of cooperation are registered on the contrary when two altruists meet (to be precise, it must be noticed that the level of cooperation in homogeneous pairs of altruistic players, in all sessions and both for low and high stake games, was the highest level in all our sample).

If we look at the heterogeneous pairs, we can see that heterogeneity has only a limited impact on the individuals' choices to choose "C". In fact, when altruists, reciprocators and selfish individuals interact in mixed pairs, we observe the same level of cooperation as in the $\gamma - \gamma$ groups. In other words, the difference in the frequencies of the choice of "C" is negligible if we compare cooperation for pairs constituted by $\alpha - \gamma$ and $\gamma - \beta$ to cooperation for pairs of altruists ($\gamma - \gamma$).[14]

Summing up these results we may conclude that, when players are aware of the social attitudes of their partners, there are significant changes in their incentives to cooperate. However, the direction of these changes is rather surprising. In fact, though we can say that – overall – information on social preferences hampers cooperation in both the low-stake and the high–stake Centipede games, the analysis of groups and individuals tell us a more complex story.

In the case of a couple of selfish individuals, cooperation decreases sharply as players do not choose to continue if they are facing a subject with similar social attitudes. On the contrary, α players are prepared to cooperate (especially in the high stake framework), when

[14] Table 7 actually shows that the highest level of cooperation was found in heterogeneous pairs of selfish and reciprocating players. However, the low number of observations we have in this case suggests caution in the interpretation of the evidence.

they are confident their partners will reciprocate – in other words, when they are facing an altruist or a reciprocator.

Altruists do not change their behaviour as much as α players do, when moving from a homogeneous pair to a heterogeneous one.

This joint effect – i.e., the constant rate of cooperation of reciprocators and altruists, both in homogeneous and heterogeneous pairs - and the significant increase of cooperation of selfish individuals - when they move from a homogeneous pairs to a mixed one - explain the higher level of cooperation we find in the heterogeneous groups.

6. Conclusions

This paper has pointed to a deeper understanding of reciprocity as the outcome of *social learning* which takes place between the two players of the Centipede. Self-interested individuals endowed with complete information about strategies and payoffs are stimulated to take the *whole* game into account at *each* decision node, and behave quite differently from the prediction of the Nash equilibrium solution. In performing the comparison between the incentive to cooperate in the presence of information on the co-players' social preferences and the incentive to cooperate in the presence of increasing payoffs in the Centipede games, we arrive at three main conclusions.

First, the level of cooperation is sensitive to the payoff structure, in as much as the increase of the payoff structure produces a sharp decrease in cooperation. This effect is stronger when players are aware of the opponents' social preferences.

The second interesting result regards the incentive to cooperate between different groups. Studying cooperation in homogeneous and heterogeneous pairs, we find that these latter are more cooperative, both in the low and in the high stake games.

In our opinion, the reason why cooperation is higher in the heterogeneous pairs is that selfish players modify their decision towards cooperation when they understand that trustworthiness could be established during the game, as they got the information that they are facing a reciprocator or an altruist. Indeed, selfish players are willing to modify their behaviour if they perceive that it is possible *and* worthwhile to trust their partners.

In the experimental literature, the debate about whether to have identical preferences or to be equally committed to compliance to a social norm – e.g., conditional cooperation – is decisive in fostering reciprocal behaviour has many contradictory aspects.

On the one hand, reciprocity is often interpreted as stemming from the sharing of the same preferences, due to the belonging to the same ethnic group (Alesina and La Ferrara, 2005). Once players know that they all have similar preferences, team reasoning develops among like-minded agents (Sugden, 1993). The internalization of social rules fosters the identification with the behaviour dictated by the group of belonging, so that an agent suffers from a loss in utility after any deviation by a member of his group from the behavioural rule he shares (Akerlof and Kranton, 2000). In experiments on a public goods game, Gatcher and Thoni (2004) found that cooperative behaviour is much more frequent in case of groups of like-minded subjects than when group composition is random.

On the other hand, a minority view argues that conditional cooperation is more often promoted by institutions devoted to facilitating the formation of social networks, even among people with different values and preferences (Habyarimana, Humphreys, Posner, and Weinstein, 2007). Our results seem to support this latter interpretation. Indeed, the information about the other player's behaviour - elicited by the strategy method and communicated to both players during the experimental session – is the "institution" which allows detecting the likely type of the other player. Both with low and high stakes, our experimental evidence also shows that reciprocating behaviour is boosted – that is, cooperation lasts longer - in the case of heterogeneous types as the mixtures of types has the effect of strengthening trustworthiness. When the player in position A is a selfish individual who has become aware that the other player is a reciprocator, the probability increases of getting a higher pay-off by sticking to node-by-node to cooperation. This suggests that the combination between a commitment to the social norm of reciprocity and self-interested behaviour sustains conditional cooperation in a multi-stage social interaction between informed individuals.

References

Akerlof G.A. and R.E. Kranton (2000), "Economics and Indentity", *Quarterly Journal of Economics*, 115: 715-53.

Alesina A. and E. La Ferrara (2005), "Ethnic Diversity and Economic Performance", *Journal of Economic Literature*, 63: 762-800.

Aumann R. (1995), "Backward Induction and Common Knowledge of Rationality", *Games and Economic Behaviour*, 8, 6-19.

Aumann R. (1996), “Reply to Binmore”, *Games and Economic Behaviour*, 17, 135-137.

Aumann R.J. (1998), “On the Centipede Game”, *Games and Economic Behavior* 23: 97-105.

Barr A., M Lindelow., J. Garcia-Montalvo, P. Serneels (2005), “Strategy Choice and Cognitive Ability in Field Experiments”, GPRG – WPS - 034.

Binmore K. (1988) “Modeling rational players: I, II”, *Economics and Philosophy*, 3: 179-214; 4, 9-55.

Benabou R. and J. Tirole (2004), “Incentives and prosocial behaviour”, Princeton University, mimeo.

Binmore K. (1996), “A Note on Backward Induction”, *Games and Economic Behaviour*, 17, 135-137.

Burlando R. M., F. Guala (2005), “Heterogeneous Agents in Public Good Experiments”, *Experimental Economics*, 8: 35-54.

Elster J. (1989), “Social Norms and Economic Theory”, *Journal of Economic Perspectives*, 3: 99-117.

Falk A., Fehr E. and U. Fischbacher (2003), “On the nature of fair behaviour”, *Economic Inquiry*, 41: 20-6.

Fehr E. and U. Fischbacher (2002), “Why Social Preferences Matter – The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives”, *Economic Journal*, C1-C33.

Fehr E. and K.M. Schmidt (1999), “A Theory of Fairness, Competition, and Cooperation”, *Quarterly Journal of Economics*, 114: 816-818.

Fehr E. and K.M. Schmidt (2006), “Experimental Evidence and New Theories”, in S-C. Kolm and J.M. Ythier (eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, Vol. I, North-Holland, Amsterdam.

Fey, M., R.D. McKelvey, and T.R. Palfrey (1996), “An Experimental Study of Constant-Sum Centipede Games”, *International Journal of Game Theory*, 25: 269-87.

Fischbacher U., S. Gachter, and E. Fehr (2001), “Are people conditionally cooperative? Evidence from a public goods experiment”, *Economic Letters*, 397-404.

Gachter S. and C. Thoni (2005), “Social learning and voluntary cooperation among like-minded people”, *Journal of the European Economic Association*, 3: 303-14.

Güth, W., S. Huck, W. Mueller (2001) “The relevance of equal split in ultimatum games”, *Games and Economic Behaviour*, 37: 161-69.

Habyarimana J., M. Humphries, D.N. Posner, and J.M. Weinstein (2007), “Why Does Ethnic Diversity Undermine Public Goods Provision?”, University of California at Los Angeles, mimeo.

Levine D. (1998), "Modeling altruism and spitefulness in experiments", *Review of Economic Dynamics*, 1, 593-622.

McKelvey R.D. and T. R. Palfrey (1992), "An Experimental Study of the Centipede Game", *Econometrica*, 60, 803-836.

Nagel R. and F.F. Tang (1998), "Experimental Results on the Centipede Game in Normal Form: an investigation on learning", *Journal of Mathematical Psychology*, 42: 356-384.

Nagel T. (1970), *The Possibility of Altruism*, Clarendon Press, Oxford.

Ponti G. (2000), "Cycles of Learning in the Centipede Game", *Games and Economic Behavior* 20: 115-141.

Rabin M. (1993), "Incorporating fairness into game theory and economics", *American Economic Review*, 83, 1281-1302.

Reny P. (1992), "Rationality in extensive-form games", *Journal of Economic Perspectives*, 6, 105-18.

Rosenthal R. (1981), "Games of Perfect Information, Predatory Pricing, and the Chain-store Paradox", *Journal of Mathematical Psychology*, 25, 92-100.

Selten R. (1967), "Die Strategiemethode zur erforschung des eingeschränkt rationalen verhaltens in rahmen eimes oligopolexperiments" in Sauerman, H. (ed) Beiträge Zur Experimentallen Wirtschaftsforschung, Tübingen: J.C.B., Mohr (Paul Siebeck).

Selten R. (1975), "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game Theory*, 4, 25-55.

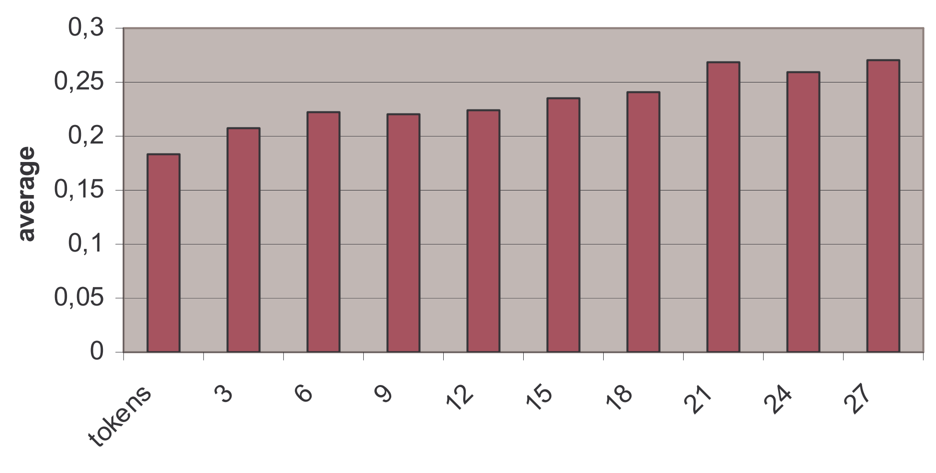

Fig. 1: Respondents' Behaviour (Session 1)

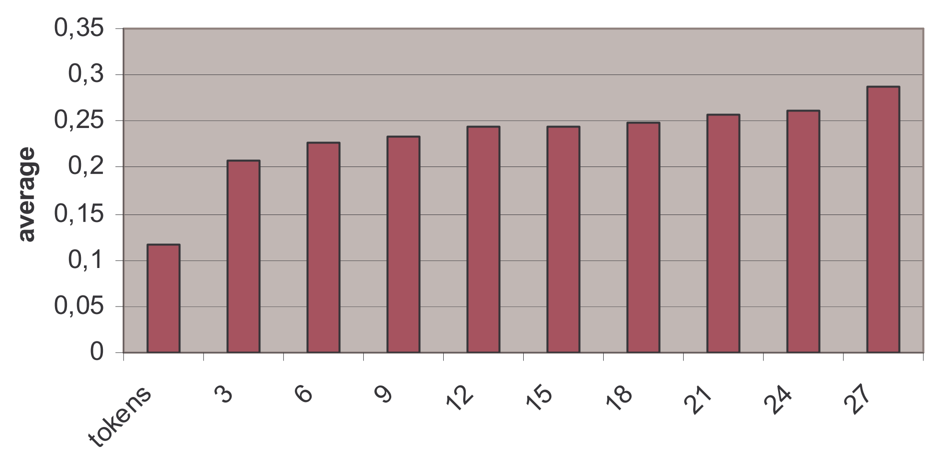

Fig. 2: Respondents' Behaviour

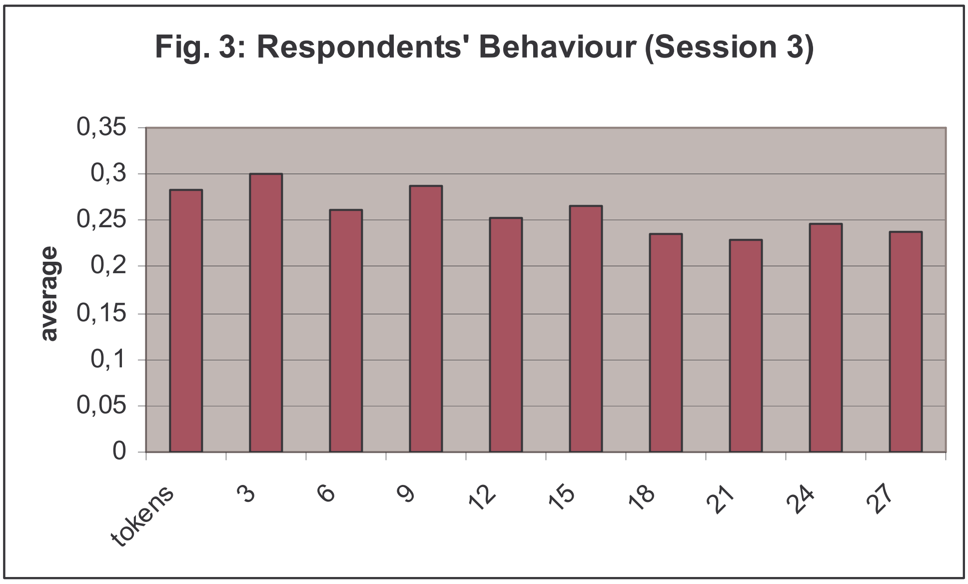
Fig. 3: Respondents' Behaviour (Session 3)
average
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
tokens
3
6
9
12
15
18
21
24
27

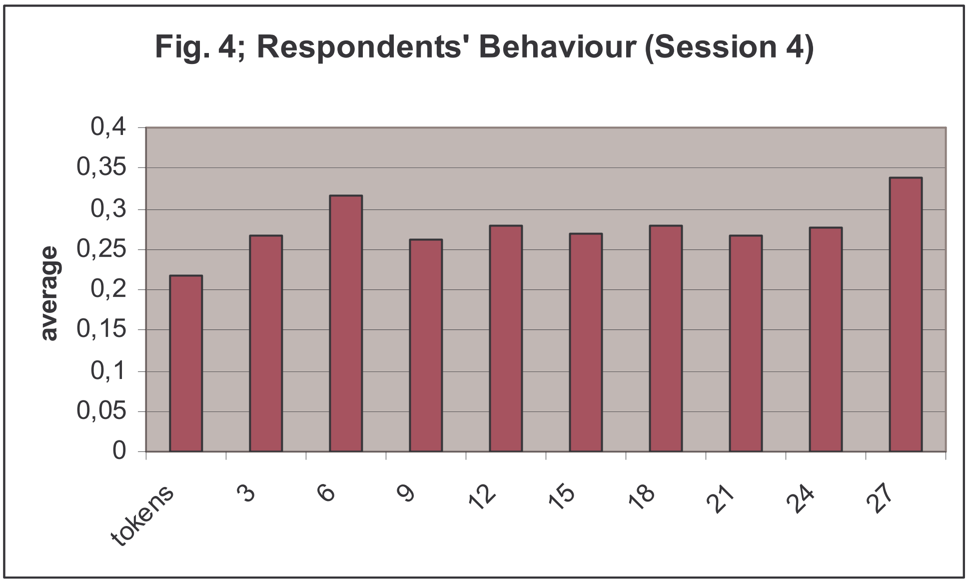
Fig. 4; Respondents' Behaviour (Session 4)
average
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
tokens
3
6
9
12
15
18
21
24
27

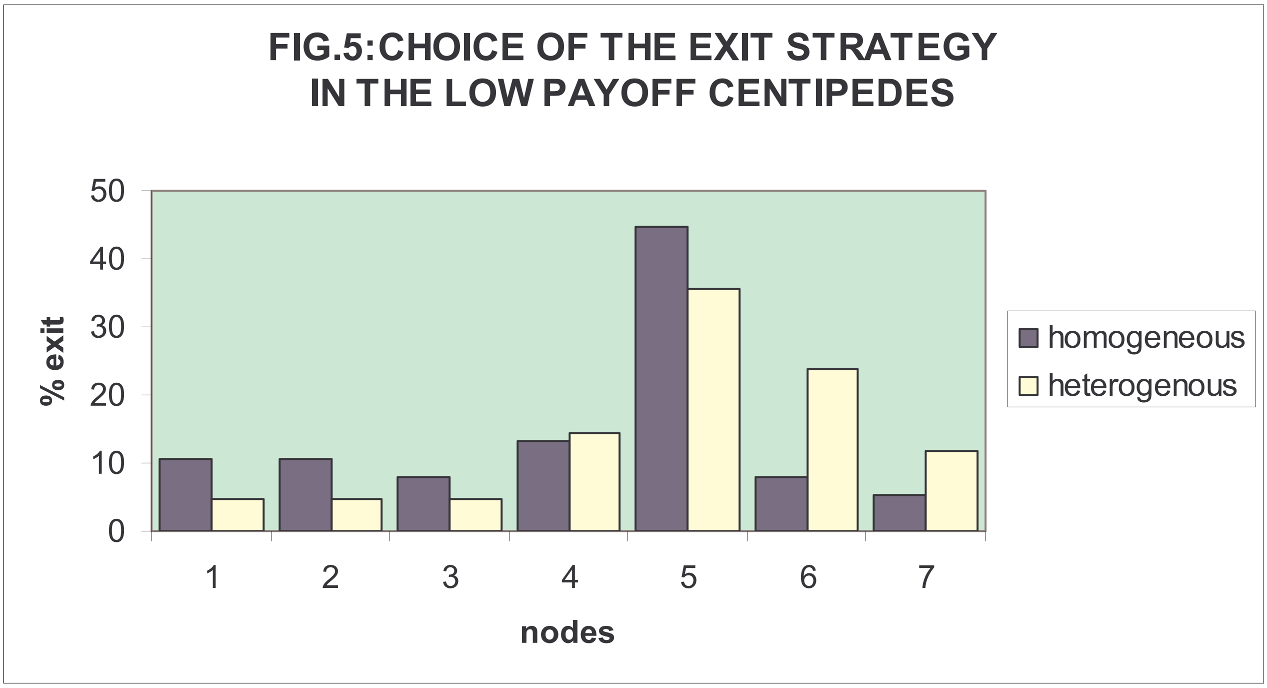
FIG.5:CHOICE OF THE EXIT STRATEGY
IN THE LOW PAYOFF CENTIPEDES
% exit
0
10
20
30
40
50
1
2
3
4
5
6
7
nodes
homogeneous
heterogenous

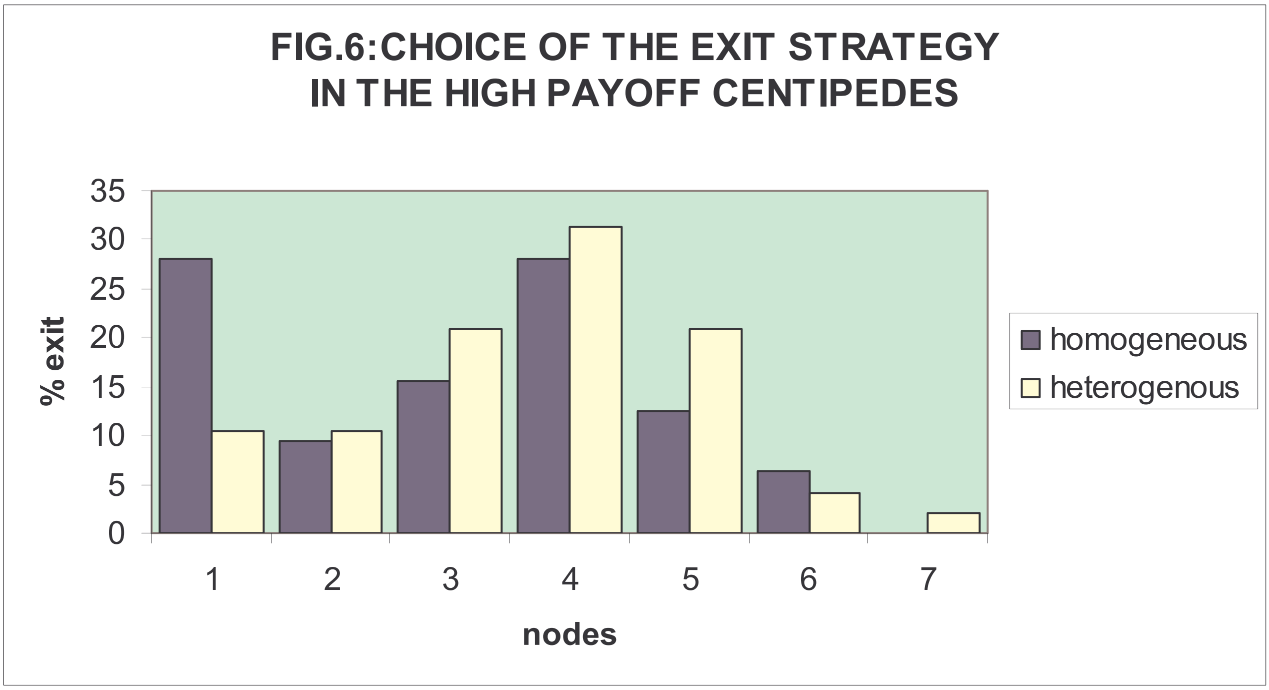
FIG.6:CHOICE OF THE EXIT STRATEGY
IN THE HIGH PAYOFF CENTIPEDES
% exit
0
5
10
15
20
25
30
35
1
2
3
4
5
6
7
nodes
homogeneous
heterogenous

TABLE 3: CHOICES OF THE EXIT STRATEGY IN HOMOGENEOUS AND HETEROGENEOUS GROUPS (Session 1)

PAIRS	11-1		12-2		13-3		14-4		15-5		16-6		17-7		18-8		19-9		20-10	
TYPE	β	α	γ	α	α	α	Α	α	α	α	γ	α	α	α	α	α	γ	γ	γ	α
EXIT (LOW)	1		5		5		5		1		3		5		5		7		5	
PAIRS	13-1		14-2		11-3		12-4		19-5		20-6		15-7		16-8		17-9		18-10	
TYPE	α	α	γ	α	β	α	Γ	α	γ	α	γ	α	α	α	γ	α	α	γ	α	α
EXIT (HIGH)	4		4		3		5		4		2		1		1		5		2	

PAIRS	1-12		2-11		3-14		4-13		5-16		6-15		7-18		8-17		9-20		10-19	
TYPE	α	γ	α	β	α	α	α	α	α	γ	α	α	α	α	α	α	γ	γ	α	γ
EXIT (LOW)	5		4		4		5		3		2		5		3		4		5	
PAIRS	1-14		2-13		3-12		4-11		5-20		6-19		7-16		8-15		9-18		10-17	
TYPE	α	α	α	α	α	γ	α	β	α	γ	α	γ	α	γ	α	α	γ	α	α	α
EXIT (HIGH)	4		1		1		3		4		3		2		1		4		1	

TABLE 4: CHOICES OF THE EXIT STRATEGY IN HOMOGENEOUS AND HETEROGENEOUS GROUPS (Session 2)

PAIRS	11-1		12-2		13-3		14-4		15-5		16-6		17-7		18-8		19-9		20-10		1-12		2-11		3-14		4-13		5-16		6-15		7-18		8-17		9-20		10-19	
TYPE	α	γ	α	α	α	β	B	α	γ	β	β	γ	α	γ	α	α	α	α	γ	γ	γ	α	α	α	β	β	α	α	γ	β	γ	γ	γ	α	α	α	α	γ	γ	α
EXIT (LOW)	6		5		7		6		5		6		2		3		5		3		6		5		1		5		5		6		6		5		6		6	
PAIRS	13-1		14-2		11-3		12-4		19-5		20-6		15-7		16-8		17-9		18-10		1-14		2-13		3-12		4-11		5-20		6-19		7-16		8-15		9-18		10-17	
TYPE	α	γ	β	α	α	β	A	α	α	β	γ	γ	γ	γ	β	α	α	α	α	γ	γ	β	α	α	β	α	α	α	β	γ	γ	α	γ	β	α	γ	α	α	γ	α
EXIT (HIGH)	4		2		4		5		5		3		4		3		2		6		2		1		1		4		4		5		3		4		4		5	

TABLE 5: CHOICES OF THE EXIT STRATEGY IN HOMOGENEOUS AND HETEROGENEOUS GROUPS (Session 3)

PAIRS	11-1		12-2		13-3		14-4		15-5		16-6		17-7		18-8		19-9		20-10	
TYPE	α	α	α	γ	α	γ	Γ	α	α	α	γ	α	γ	α	α	γ	α	γ	α	α
EXIT (LOW)	4		6		5		6		5		5		7		7		5		4	
PAIRS	13-1		14-2		11-3		12-4		19-5		20-6		15-7		16-8		17-9		18-10	
TYPE	α	α	γ	γ	α	γ	Α	α	α	α	α	α	α	α	γ	γ	γ	γ	α	α
EXIT (HIGH)	5		3		4		4		3		1		1		3		6		4	

PAIRS	1-12		2-11		3-14		4-13		5-16		6-15		7-18		8-17		9-20		10-19	
TYPE	α	α	γ	α	γ	γ	α	α	α	γ	α	α	α	α	γ	γ	γ	α	α	α
EXIT (LOW)	6		4		5		6		4		2		7		1		2		5	
PAIRS	1-14		2-13		3-12		4-11		5-20		6-19		7-16		8-15		9-18		10-17	
TYPE	α	γ	γ	α	γ	α	α	α	α	α	α	α	α	γ	γ	α	γ	α	α	γ
EXIT (HIGH)	4		5		5		4		2		1		4		3		6		3	

TABLE 6: CHOICES OF THE EXIT STRATEGY IN HOMOGENEOUS AND HETEROGENEOUS GROUPS (Session 4)

PAIRS	11-1		12-2		13-3		14-4		15-5		16-6		17-7		18-8		19-9		20-10	
TYPE	α	α	γ	α	γ	γ	Γ	β	γ	α	β	β	α	γ	α	γ	α	γ	γ	α
EXIT (LOW)	4		5		5		4		5		1		7		4		5		1	
PAIRS	13-1		14-2		11-3		12-4		19-5		20-6		15-7		16-8		17-9		18-10	
TYPE	γ	α	γ	β	α	γ	Γ	β	α	α	γ	β	γ	γ	β	γ	α	γ	α	α
EXIT (HIGH)	4		5		5		4		5		1		5		5		2		3	

PAIRS	1-12		2-11		3-14		4-13		5-16		6-15		7-18		8-17		9-20		10-19	
TYPE	α	γ	α	α	γ	γ	β	γ	α	β	β	γ	γ	α	γ	α	γ	γ	α	α
EXIT (LOW)	5		5		5		6		7		5		4		5		2		2	
PAIRS	1-14		2-13		3-12		4-11		5-20		6-19		7-16		8-15		9-18		10-17	
TYPE	α	γ	α	γ	γ	γ	β	α	α	γ	β	α	α	β	γ	γ	γ	α	α	α
EXIT (HIGH)	4		3		4		3		7		1		3		6		4		1	

TABLE 7: FREQUENCIES OF THE "CONTINUE" STRATEGY PER TYPE AND GROUPS

TYPES	α - α	β - β	γ - γ	α - β	α - γ	γ - β
LOW	0.80	---	0.83	0.89	0.83	0.83
HIGH	0.74	---	0.77	0.70	0.78	0.80

....