QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

MAURO CAMINATI
ALESSANDRO INNOCENTI
ROBERTO RICCIUTI

Drift Effect Timing without Observability: Experimental evidence

n. 405 – Ottobre 2003

Abstract - We provide experimental evidence to Binmore and Samuelson's (1999) insights for modeling the learning process through which equilibrium is selected. They proposed the concept of drift to describe the effect of perturbations on the dynamic process leading to equilibrium in evolutionary games with boundedly rational agents. We test within a random matched population two different versions of the Dalek game where the forward induction equilibrium weakly iterately dominates the other Nash equilibrium in pure strategies. We also assume that the first mover makes her decision first ("timing") but the second mover is not informed of the first mover's choice ("lack of observability"). Both players are informed of their position in the sequence and of the fact that the second player will decide without knowing the decision of the first player. If the actual observed choices are only those made by other players in previous interactions, the role played by forward induction is replaced with the learning process taking place within the population. Our results support Binmore and Samuelson's model because the frequency of the forward induction outcome is payoff-sensitive: it strongly increases when we impose a slight change in the payoffs that does not change equilibrium predictions. This evidence reinforces the evolutionary nature of the drift effect.

Keywords: evolutionary games, experiments, drift, forward induction, order of play.

Jel Classification: C72, C91

We thank Francesco Lomagistro for writing the experimental software and for research assistance. Research support was provided from Miur and University of Siena. Previous versions were presented at the 7th CEEL Workshop in Experimental Economics (Trento, May 2003) and at the ESA European Conference (Erfurt, September 2003).

Mauro Caminati, Dipartimento di Economia Politica, Università degli Studi di Siena
Alessandro Innocenti, Dipartimento di Economia Politica, Università degli Studi di Siena
Roberto Ricciuti, Department of Economics, Royal Holloway, University of London

1. Introduction

The assumption of strong rationality characterizing game theory has been widely discussed and criticized both by theorists and experimentalists. Most criticisms have been targeted to the toolbox of solution concepts that has been modified and extended to improve the applicability of game theory to actual choices. A set of equilibrium refinements is based on the general idea that beliefs on off-the-equilibrium-path information set can affect players' strategy. As a result, Selten's trembling hand perfection, Kreps and Wilson's sequential equilibrium, Myerson's proper equilibrium, and Kohlberg and Mertens's forward induction have also become established objects of experimental work.

A theory of equilibrium selection follows the logic of backward-induction that results from coupling Harsany and Selten (1988) notions of subgame consistency ("Play in a subgame is independent of the subgame position in the larger game") and of truncation consistency ("Replacing a game with its equilibrium pay-offs does not affect play elsewhere in the game").[1] In games of imperfect information the possibility of multiple equilibria in a subgame may require a selection between these equilibria, before the logic of backward induction can be applied.

Another class of equilibrium refinements follows from the general logic of *forward induction* stating that in games of imperfect information players at an information set can deduce their plausible beliefs from the information disclosed by previous choices in the game. Forward induction applies in situations where earlier players' actions can reveal private information to later players observing those actions. But what does it happen if observability is removed but timing does not change? This question turns out significant in an experimental setting where actual behavior is tested in repeated games to draw some insights on the dynamic process taking the players to equilibrium. If the actual observed choices are only those made by other players in previous interactions, the role played by forward induction is replaced with the learning process taking place within the population. Our experiment is focused on this issue that we consider as relevant for the analysis of equilibrium convergence in evolutionary games. Specifically, this paper intends to provide experimental evidence for Binmore and Samuelson's (1999) model of the learning process through which an equilibrium is selected. They proposed the concept of drift to describe the effect of perturbations in evolutionary games with boundedly rational agents. Our experimental design makes the assumption of timing without observability to test the relevance of the drift effect within a random matched population interacting in an evolutionary setting.

The paper is organized as follows. Section 2 discusses the theoretical background and the relationships with previous experimental literature. Section 3 describes the experimental design. Results are presented in Section 4. Section 5 draws some conclusions and suggests some future experiments.

2. THEORETICAL AND EXPERIMENTAL BACKGROUND

In games of imperfect information where a solution may possibly emerge as a result of one of the Nash equilibrium refinements, play may depend on seemingly irrelevant features of the game which make one or the other selection notion more prominent. In fact, in iterated versions of the game, these apparently irrelevant features may direct players' adaptive learning towards different selection criteria.

Repeated interaction among players as a way to study the stability properties of Nash equilibrium was the object of Binmore and Samuelson's (1999) drift effect. A strong motivation for referring to this theory on the relation between learning and equilibrium selection is that it claims to be more consistent than other theories (Young, 1993; Kandori, Mailath and Rob, 1993), at least in the absence of 'local interactions' (Ellison, 1993), with the time horizon of the processes that are more likely to be observed in the laboratory.[2]

Consistently with Binmore and Samuelson's notion of drift in games, the out-of-equilibrium states that are observed in the repeated actual play of a game can be modelled through the dynamic coupling of the following factors:

(i) Initial conditions corresponding to the selection by agents in the subpopulation of players of different strategies sustained by prior beliefs on opponent's behaviour.

---

[1] Binmore et al. (2002, 53).

[2] Roth and Erev (1995) report experiments supporting the intuition that for some classes of games the very-long-term predictions may not be observed. See also Ponti (2000) which models and tests learning processes in the centipede game.

(ii) Population learning resulting from the selection by each sub-population of players of the strategies that are relatively more profitable, given the available information on the actual play of the game.

(iii) Drift, namely, expected, sub-population-average 'mistake' in the given game situation.

(iv) Residual stochastic perturbations.

According to Binmore and Samuelson, the drift component arises as a result of a misreading of the game induced by un-careful scrutiny that causes confusion between the actual conditions of the game and those referring to similar games previously experienced in real life. To the extent that confusion arises as a result of un-careful scrutiny, the prediction is that the size of the drift component will be inversely related to the potential pay-off consequences of a mistake.

In our view, population-average behaviour, not strictly conforming to the rule (ii) above, may originate not only from confusion with similar game situations causing 'mistakes', but also from the reliability assessment attached to different modes of reasoning, including alternative selection arguments. We expect that apparently irrelevant features of the pay-off structure of the game influence the reliability in question. These 'irrelevant' features may therefore reveal themselves not only in that they can place the initial conditions of play (i) in the basins of different dynamic equilibria, but also in the selection of a different perturbation component. For instance, a small fraction of a sub-population may play a fixed strategy, irrespective of the way in which the frequency distribution on strategies in the population evolves during the repetition of the game.

The theory of drift proposed by Binmore and Samuelson (1999) predicts that under suitable conditions the dynamics induced by the coupling of *(ii)* and *(iii)* can 'stabilize' an equilibrium w^* belonging to a set of Nash-equilibrium components E that are not asymptotically stable in the learning dynamics induced by *(ii)*. The reason for this lack of dynamic stability is that at an unreached information set h a player is indifferent between the actions that are available at h. The suitable conditions refer to the 'compatibility' between drift and E. In particular, the action distribution prescribed by drift at h must belong to the relative interior of the set E_h of action distributions at h supporting some state in E as a Nash equilibrium. The larger E_h, the higher the probability that drift is 'compatible' with E.

To test this prediction Binmore and Samuelson proposed the modified version of the Dalek game (Figure 1) as the object of a possible experiment.[3] The game has a forward induction solution that is also obtainable for iterated weak dominance (Play-Red, Left)[4] and another sub-game perfect equilibrium in pure strategies (Exit, Right) as long as x is lower than 10. There is also a component E of Nash equilibria in which agents in population 1 play Exit and agents in population 2 play Right with probability higher than p. Binmore and Samuelson suggested that the frequency of the outcome (Play-Red, Left) would be lower when the payoff x obtained by player 1 with (Play-Black, Left) is relatively small (e.g., equal to 1), than when it is relatively large (e.g., equal to 6) because the size of the set of frequency distributions on Player-2 choices supporting some state in E as a Nash equilibrium is inversely related to the value of x, thus affecting the probability that drift 'is compatible with' E. The complication arising in this context and that must be considered in the evaluation of the experimental evidence, is that the slight change of pay-off across the two games affecting the critical value of p, may affect factors other than the size of E. It could induce a change in the subjects' behavior, that can be considered as relevant either because it modifies the initial conditions of the game (as admitted by Binmore and Samuelson), or because it modifies the size and specification of the drift component (as we suggested), or both.

In terms of the notation used above, in the proposed game the solution (Exit-Right) is a state in E. For ease of exposition, the play of the game will be said to correspond to this solution if the population-distribution of play converges to an action in E. One can presumably expect that the frequency with which the play of the game corresponds to the two solutions is influenced by the pay-off modifications leaving the sub-game perfect equilibria unchanged, but affecting: *(a)* the size of the set E_h; *(b)* the agents' prior beliefs, hence the initial conditions of the repeated play; *(c)* the size and specification of the drift component.

Our experiment does not completely adhere with Binmore and Samuelson's proposal. We test actual behavior in the two games represented in Figures 2a and 2b (extensive form) and in Figure 3 (normal form).[5] To analyze the effect of repeated interaction in a random-matched population we assume that the actual observed choices by players 2 are only those made by other sub-population of players in previous interactions. This assumption rules out

---

[3] The Dalek Game was originally proposed by Kohlberg and Mertens (1986) and discussed by Binmore (1987-1988) to illustrate the solution concept of forward induction.

[4] (a) For player 1 Exit dominates Play-Black; (b) by deleting Play-Black, for player 2 Left weakly dominates Right; (c) with regard to Left, for player 1 Play-Red dominates Exit.

[5] The normal form of the game is represented here for clarity of exposition but it was not showed to the experimental subjects.

interactions between members of the same sub-population. In particular, players had no information on the frequency distribution of play within their own sub-population.[6] More significantly, our variant also assumes that player 2 does not know the actual choices of player 1 at the time they are taken even if she knows that player 1 chooses before her. At the end of each repetition both players are informed of the particular choice made by their counterpart and of the distribution of choices made in the other-player sub-population.[7] This assumption has two other advantages. First, players 2 take their first decision conditioning their behavior on the pay-offs and not as a reaction to the choice of the leaders. Second, if player 1 chooses the outside option we are able to observe the actual choice of player 2, which in fact is not taken.

This design allows us to inquire what happens to the equilibrium if observability is removed but timing does not change. Game theorists that typically assume that an unknown event does not affect subsequent choices even if it happened before neglect this issue. This long-standing view[8] follows from the idea that priority of information includes indeed priority in time but not vice versa, and this makes the former the basic way of characterizing strategies.

Experimental work gives evidence for the opposite idea that timing of unobserved moves could matter.[9] In a seminal work, Rapoport (1997) showed how players, when they choose in a predetermined sequence without knowing previous other players' choices, behave differently from the way they do when all players move simultaneously. The main finding of Rapoport's experiment is that subjects are inclined to select the strategy that gives the first-mover his preferred outcome. Güth et al. (1998) provided further experimental evidence that this attitude could increase the probability of disequilibrium play, although this effect can be weakened by the existence of an outcome of the game that is clearly perceived as fair. More recently, Weber and Camerer (2001) and Muller and Sadanand (2003) studied the effect of the order of play in laboratory experiments. Weber and Camerer found that the timing of moves without observability affects behavior in experiments on coordination games. They consider three person "weak link" games such as the stag hunt, where efficient equilibria are dependent on the sequence of moves. Their results confirms that outcomes in condition of timing

[6] In this way, although we intend to make the learning process taking place within the population the very object of our test, we abstract from the way learning may be affected by fashion and convention *within a sub-population.* For a consideration of these factors in a model of drift see Uriarte (2003).
[7] The same applies to sub-population 1.
[8] See Von Neumann and Morgenstern (1947, 51).
[9] See Weber and Camerer (2001) for a survey.

without observability are closer to games with the fully observed condition than to simultaneously played games. Muller and Sadanand’s work confirmed that the mere knowledge of the order of play affects players’ actions in very simple two-person games. They found that the effect of timing without observability depends mainly on presentation form: it is more evident if games are presented in normal-form than in extensive form. Interestingly enough, extensive forms presentation seems to induce a sequential mode of reasoning even with simultaneously played game.

A theoretical explanation of this evidence can be based on the way the mere order of play influences players’ reasoning process. A proposal in that direction came from Amershi et al. (1992) that coined the notion of Manipulated Nash equilibrium. They inquired the applicability of forward induction in signaling games where the leader makes her choice first and the follower knows the sequence of decisions without being informed of the leader's choice. In this setting behavior could be predicted as if the follower could observe the earlier player’s choice by assuming that the leader asks herself what the follower would believe about leader’s previous unobservable choices and what the follower’s best response would be. According to this reasoning the leader would choose the best course of action on the basis of the anticipated behavior that she deduced about the follower. If the follower expect the leader to follow this reasoning and to act consequently, she would make her best choice given leader’s best choice. In this way the theory predicts that earlier players can induce later players to make choices that are most favorable to the former. The selected equilibrium would be that predicted in the case of fully observed play rather than in the case of simultaneous unobserved play.

The concept of Manipulated Nash equilibrium is replicated and extended by the theory of virtual observability proposed by Camerer et al. (1996). It predicts that in games of imperfect information in which previous moves are unobservable players will behave as if they could observe previous moves as long as there is a subgame perfect equilibrium to the game with observable moves. Virtual observability implies that the knowledge of timing with unobservable moves triggers a chain of reasoning articulated in five steps:

(i) The simultaneous game is translated in a game of imperfect information in which all earlier moves are unobservable.

(ii) Given the timing of the moves, all previous moves are considered as they were observable.

(iii) The subgame-perfect equilibria of the game with observable moves is selected.

(iv) If the selected equilibrium is also a Nash equilibrium of the original simultaneous game, it is selected as the equilibrium to the game.

(v) If not, timing with no observable moves does not influence play.

Later players' reasoning process would work as if they "break" information sets of the game with unobservable moves in order to apply subgame perfection and select the same Nash equilibrium of the game with observable moves.[10]

Following the main concept pointed out by this literature, in the two games proposed by our experiment the mere order of play may act as a coordination device allowing players 2 to focus into the equilibrium that is preferred by Player 1. The same prediction would follow from the forward induction argument if Player 2 had information on Player 1 previous choice. In our design Binmore and Samuelson's prediction of convergence to the outcome (Exit, Right) in Game A and to the outcome (Play-Red, Left) in Game B can be predicated from the effect of pay-off modification on: *(a)* the size of the component E of equilibria, in conformity with Binmore and Samuelson's theory of drift; *(b)* the salience of the virtual-observability versus the backward-induction selection argument affecting both the initial conditions of play and the inclination to change one's strategy in conformity with the adaptive rule *(ii)* above.

We conjecture that in the given game situation players have a possibly confused perception of the different selection arguments and are truly in doubt concerning the force of such arguments. In conditions of true uncertainty, the perceived salience of a coordination device in the mind of a player -repeatedly meeting an opponent randomly sorted out with uniform probability from a sub-population of players- is affected by the size of the set of frequency-distributions on the opponent's strategies capable of causing regret on the part of that agent for her choice of the coordination device. Notice that the size of the set referred to above is independent of the actual distribution of play in the previous repetition of the game, hence it is a factor which may add a fixed perturbation component throughout the learning process.

Our experiment is indeed a variation of Balkenborg's (1994) experiment that tested a version of the Dalek game with x equal to 0. He found that by making the standard

[10] It is useful to recall that timing without observability can be related to Schelling's notion of salience in co-ordination games. Timing would make an equilibrium solution more salient because it could brings asymmetry into the game It is also worth adding that in some game situations the same forward-induction solution

assumption of observability the forward induction option was virtually always discarded. In game A of our test x is equal to 1 and we presume to obtain similar results to Balkenborg's experiment. At the same time we expect to find that in Game B the increase of x from 1 to 6 will be effective in selecting the equilibrium (Play-Red, Left) even if the only choices observed by players 2 are those made by players 1 in the previous periods. Strictly speaking, our design attempts at studying the dynamic process induced by the adaptive mode of behavior characteristic of evolutionary learning in a context where players are truly uncertain concerning the force of the rules of behavior following from careful reasoning in single game interaction.

3. LABORATORY PROCEDURES

We have designed a baseline treatment and two variations to check for the robustness of the results. Each session was divided in two sub-sessions, A and B, in which we submitted to subjects respectively games A and B represented in Figures 2a and 2b. The working of the experiment is as follows. There are 12 players that are randomly divided in two groups, 1 and 2. Once each player is assigned to one group, she belongs to it for the whole duration of the experiment. Each member of one group is randomly matched with a member of the other group. At each repetition of the game each player of one type has a uniform probability of being matched with a player of the other type. These features are publicly announced. Players 1 (which in the instructions are called "blue" and in the remainder of the text are also called leaders) start the game by choosing either Exit or Play. In the first case they obtain a prize of 7, whereas the opponent gets 5. Players 2 (defined in the instruction as yellow, and also called followers henceforth) make a second move by choosing either Right or Left without knowing whether his counterpart has decided to exit or to play. If player 1 has decided to play, at this stage he has to choose between Red and Black. It is important to emphasize the information players are given. At the end of the experiment, when they receive the outcome for that period, each player is informed about her and opponent's decisions. In addition, leaders are informed about the number of followers who have chosen Right and Left. In turn, players 2

---

predicated by the theory of virtual observability is also obtained through iterated elimination of weakly dominated strategies in the normal form of the game, a context where timing is clearly irrelevant.

are informed how many players 1 have chosen Exit and Play, and how many have chosen Red and Black out of those who have chosen to play.

The experiment starts with the game A and after 40 repetitions we turn to the game B. When the instructions are distributed and read we only give the sheet representing the game A, to prevent confusion and strategic play between the two treatments. At the beginning of sub-session B we have distributed the relevant decision tree. Only before sub-session A, a five-period trial was done to get subjects acquainted with the experiment and the software. The earnings opportunities in this experiment are very asymmetric between the two groups, therefore the show-up fee was worth € 3 for the players 1 and € 7 for players 2. Doing so we aimed at avoiding regret and related misbehavior from players 2 subjects that are clearly disadvantaged by the payoffs structure. The variable part of the earning is obtained by multiplying the number of points obtained in the 80 periods for 2 eurocents. The exchange rate was announced to the subjects before the session started. The average payment for this part was about € 12 for subjects in role 1 and € 8 for subjects in role 2.

To check for the robustness of the outcome of the baseline treatment we ran another treatment inverting the order between sub-sessions A and B, leaving all other features unchanged.

A second check was done to controlling for the risk attitude of the subjects. To implement it, we have conducted a third treatment using a modified version of the binary lottery proposed by Roth and Malauf (1979). We gave a fixed payment of € 10 to each player and an additional € 10 prize for a person in each group based on a lottery on the relative performance of each player.

Table 1 sketches the experimental design by session. We conducted five sessions of the baseline,[11] and two sessions for each control treatment during the academic year 2002/03. The subjects were second-year undergraduate students in economics at Siena University. Each session lasted approximately an hour. During the sessions the subjects were seated at computer terminals in separate seats to prevent direct communication or visual contact among them. The experiment was computerized using the software Z-tree (Fishbacher, 1999).

[11] Session 5 of the baseline was made up by 10 subjects divided in two groups of 5 each, because two students did not show up.

4. RESULTS

In the light of the discussion of Binmore and Samuelson (1999), we tested the following hypothesis:

In treatment A the outcome (*Exit, Right*) *is more likely to occur than* (*Play-Red, Left*), *whereas the opposite is expected in treatment B.*

We start our analysis by looking at the behavior of players 1 and 2 and the relevant outcomes in the baseline for game A, then we consider these features for game B. Furthermore we analyze the evidence for the two other treatments (risk control and reverse order). To describe the learning process in these games we divide the 40 repetitions of each game in 8 groups of 5. To simplify the discussion, outcomes are indicated as follows: A = (Exit, Left), B = (Exit, Right), C = (Play-Red, Left), D = (Play-Red, Right), E = (Play-Black, Left), F = (Play-Black, Right).

Figure 4a shows that usually less than 50% of players decided to exit in all the 8 groups of periods considered here, with an average of 46.3 (see Table 2) and ranging from 41.4% to 50.3%, the latter being the outcome of the last group of 5 repetitions, which can therefore be related to an end-game effect. While the decision to exit does not change much, the composition of the choices resulted from playing the game shows stronger dynamics. In the first group the strategy Red is taken in the 37.2% of cases, whereas the strategy Black is taken 17.9% of cases (Table 2). In the second group they are respectively 47.6% and 11.0%. On average they are 45.1% and 8.6%, respectively. The Black strategy strongly decays and then stabilize, while there are players that switch from time to time from Exit to Red and vice versa. The majority of followers starts playing Left, then there is a reversion of this strategy in subsequent plays, and from the third group on the preferred strategy is Left (Figure 6b). The average for both strategies is 53.5% and 46.5% (Table 2), respectively, and also in this case and end-game effect may be apparent. Figure 4c shows that the forward induction equilibrium (C) does not emerge as the actual consistent solution of the game as it happens in Balkenborg's experiment. On average it amounts for 30.6% of the cases (Table 2), starting from 20% and then increasing in the first few groups of periods stabilizing in a range between 30.3% and 36.6%. Nonetheless, it is the most likely outcome of game A. The more likely outcome according to our Hypothesis (B), is the second most attained result (though it is the most played one in some of the first periods of the game), with an average of 26.4% and a rather low volatility. Other out-of-equilibrium outcomes still amount for about 40% of the results, with an increase in the plays of A and a decrease for D, E, and F. However, in aggregate learning to eliminate clearly dominated strategies does not seem to take place in a clear way.

The picture is quite different when we turn to game B. Leaders start playing the Red strategy soon at a very high rate and then tend to converge to 100% (Figure 5a). In fact, this strategy is played on average 95.3% of cases (Table 2). The other two strategies are rarely played, and in many instances Black is never played. Players 2 show a different pattern (Figure 5b). In this case too, the Left strategy accounts for a large number of plays in the first group, but here a learning process appears to be slowly taking place, and does not go very close to 100%. As a result, the average for Left and Right are 86.5% and 13.5%, respectively. Finally, C accounts for the large majority of the plays (Figure 5c). It starts just above 60% and increases over time with and average of 82.7% and a share in the last four periods of about 90%. Out-of-equilibrium outcomes account for a very low share of the plays. Outcomes A, B, E, and F are played very few times, consistently among the group of periods, whereas outcome D sees its share reduced for the benefit of C. Indeed, the two lines are almost symmetric.

We can also state that our results support the hypothesis that the change in outcomes induced by the change in pay-off is partly explained by their influence on the initial conditions of the learning process, that is, on the agents' initial beliefs. Indeed, comparing the distribution of strategies played by a sub-population in the first repetitions of the games A and B, respectively, we notice that they are quite different. For example in the first 5 plays of game A Exit is played 44.8% of times, Red 37.2%, and Black 17.9%, while in game B the frequencies are 7.6%, 86.9%, and 5.5%, respectively (Table 2). In turn, followers play Left 45.5% and Right 54.5% in game A, and 73.8% and 26.2%, respectively in game B.

Results do change for game A when we induce risk neutrality, strengthening our hypothesis. Table 3 shows that Exit is played more than in the case without risk control. It averages to 67.9% of plays. Red and Black are therefore much less chosen. For players 2 there is more difference between Left and Right than before: relevant averages are 49.4% and 50.6%. In the first repetitions (Figure 6a) leaders prefer Black to Red as in the other experiment but much more, then there is a deep fall in the use of the strategy Black. As a result, Figure 6c reveals a rather different picture: B is the most likely outcome and it is reached in all but one of the groups of periods (the last one, when the forward induction outcome slightly prevails). However, with the exception of the first group of periods, it only accounts for a relative majority of the outcomes (38.1% on average), while the forward induction equilibrium is obtained in 17.7% of cases. One should note (Table 3) that A emerges as the second most attained outcome (29.8%), and that initial outcomes are not reversed but a much wider choices dispersion is observed.

In contrast, in game B with risk control does not show significant differences with respect to the other one. Figure 7a reveals a slightly slower convergence to 100% of plays by players 1 to the Red strategy and this level is reached in two groups of periods differently from the other experiment. Players 2 tend to reach the 100% of plays for the Left strategy faster than without risk control (mean 92.1%). Finally, outcome C does emerge as the equilibrium of the game in the vast majority of the periods, averaging 85.8% of total plays (Figure 7c and Table 3). Note that Table 3 reports reversion of outcome in this experiment: the initial outcome being Play-Red/Left (D).

For player 1 the choice Play is riskier game A than in game B. Therefore, the critical value of the probability that makes him indifferent between Exit and Play is closer to 1 in game A, which makes Exit more attractive. The difference within game A between the one with induced risk-neutrality and the other without control can be attributed to two features. The first is straightforward and appeals to the circumstance that players recruited in these sessions are in fact risk-lovers. The second is based on a more careful analysis of the plays in each of the five sessions of the baseline. Indeed, in two sessions there are a larger number of players 2 than in the others that consistently choose Right. Therefore, players 1 find more attractive to secure the pay-off associated with Exit when facing more aggressive players 2.

Another control involved reversing the order in which the two games were played: first game B and then game A. One may believe that the equilibrium selection process in each game could affected by the order in which they are presented. For game A results are fairly in line with those of the baseline. As Table 4 shows, on average players exit 50% of times, which is slightly higher than in the baseline, though the pattern is more volatile (Figure 8a). Another difference concerns the Black strategy, which is chosen much less (4% versus 7.6%) and indeed it is never played in the last 10 periods. Figure 8b shows a higher variability for players 2, which on average choose Left 48.1% and Right 51.9% of times, in slight contrast with the baseline. The picture resulting from outcomes (Figure 8c) shows that outcomes B (Exit, Right) emerges consistently as the most reached end-node of the game (on average

30%). As it can be expected given the variability of strategies of the two players, the pattern is much more volatile. Table 4 illustrates that there are small differences between initial and average plays, the most notable concerns the emerging outcome, which is Play-Red/Left (C) in the first period and then it is overcome by Exit/Right (B).

The picture emerging for game B is more in line with previous findings, though we do not observe the virtual disappearance of some strategies from the actual choices of both types of players. For example, players 1 choose Red 74% of times against 95.3% in the baseline, and while Black is played a few times and tend to disappear at the end of the game, Exit is played on average 21.3% of times with a decreasing pattern (Figure 9a). The pattern is interesting: in the baseline players start choosing Red at a very high rate (about 80%) and then there is an upward trend that brings the share of this choice at about 100%. Here the starting level is quite lower (about 60%) to end at about 85%. Learning seems to explain these results. In the baseline subjects had already played game A for 40 periods, therefore they were quite acquainted with the game, and soon noticed the change in parameter, and changed their choices accordingly. In the reverse order treatment game B was the first to be played so subjects had to learn how to play in this game. Yet, the learning dynamics in this game is impressive. We can argue that two learning process take place in these sessions: one within each experiment and another between experiments. The same happens for players 2, which on average choose Left 69.6% of times (versus 85.5%) and Right 30.4% (versus 13.5%), but with a limited tendency to increase/decrease the number of choices, respectively. The outcome C (Play-Red, Left) is played 53.8% of times against 82.7%, starting from 38.3% and ending at 68.3% of times.[12] All other non-equilibrium nodes are reached a much higher number of times, with the prevalence of D (Play-Red, Right). Table 4 reports an impressive similarity between initial and average plays for all the strategies and outcomes.

A formal investigation of the relationship between the choice of exiting from the game (defined as state 0) versus the choice to play (state 1) for Players 1 as a response to the aggregate choices of Players 2 can be analyzed using a grouped logit estimation model in which the choice between Exit and Enter[13] (subsequently playing either Red or Black) is a function of a constant and the ratio of Left choices to Left plus Right choices (LRATIO) in the previous period. This is meant to capture how Leaders react in the next period to the

---

[12] This pattern cannot be seen in Table 4 where the initial and average results for outcome C are very close. It appears that agents first started playing C, then drastically decreased the number of this choice and than increased it, as noted above.

[13] We interchangeably use the phrases "play" and "enter into the game".

composition of choices by Followers, information that they can observe, since it is given in a summary window at the end of each period.

$$\Pr(\text{Enter}) = \alpha + \beta \text{LRATIO}_{t-1} + \varepsilon_t. \tag{1}$$

We expect that as long as the number of Players 2 who chooses Left increases, Players 1 are more likely to enter into the game because in this case they can obtain a higher payoff. Therefore β should be positive in both experiments.

Table 5 reports the results in the baseline treatment for both games corroborate this hypothesis: both coefficients of the Left choice ratio enter significantly in the two regressions, with the expected sign. The coefficient of the constant is negative for game A, indicating a tendency to exit from the game, which we have already noted in previous descriptive analysis. For game B, the constant is positive – in turn confirming previous results - but borderline insignificant. The derivative of probabilities calculated at sample frequencies for game A for LRATIO_{t-1} at state 1 is 0.693, while for game B it amounts to 0.087.[14] We interpret this difference noting that in game A subjects are not likely to enter into the game but an increase in the share of Left plays largely increases their attitude to enter. In contrast, in game B players 1 are already inclined to enter because of the structure of the pay-off, and therefore an increase in LRATIO_{t-1} has a tiny effect on their attitude.

Another econometric estimate concerns the learning process. We divide each 40-period games in two sub-periods: the first one includes periods from 1 to 20 and the second one goes from 21 through 40. With the dummy variable PERIODUM we identify the latter first sub-period with 0 and the second with 1. Following the same method previously applied, the estimated model is the following:

$$\Pr(\text{Enter}) = \alpha + \beta \text{PERIODUM} + \varepsilon_t. \tag{2}$$

We expect that over time players 1 should learn to exit from the game (therefore β should be negative) in Game A, whereas the same players will learn to enter into the game B (β be positive). In Game A the sign of β is negative as expected but is not significant. In

[14] In Table 5 no test for the significance of the overall regression is given because the relevant Chi-squared statistic is calculated only on the non-constant regressor, therefore in this case it degenerates to the squared t-statistic of LRATIO_{t-1}.

Game B PERIODUM enters significantly with the expected sign. Learning appears to take place more in Game A than Game B and this circumstance is consistent with the patterns in Figures 4a and 5a.

Finally, it is interesting to see whether some subjects have been stacked on a single strategy for the whole duration of the experiment. In the baseline game A, this has not happened for players 1 (a partial exception being a subject that played Exit 39 times and Play-Red in the remaining case), but for players 2 this was the case for 6 subjects (3 choosing Right and 3 choosing Left). For the game B, 9 leaders picked Play-Red all the times (plus 7 playing this strategy 39 times), and 7 followers always selected Left (plus 2 choosing this strategy 39 times).[15] In game A of the risk-control treatment, 2 players 1 always opted to exit, while between the players of the other group 1 choose Left and another Right 40 times. In game B, 4 players 1 always selected Play-Red (and other 2 did this 39 times) and 4 players 2 always played Left (and another one picked this strategy 39 times). In game A of the reverse order treatment, only one player 1 has always chosen to exit, and two players 2 have played Right for the whole duration of the game (but three picked 39 times Left). In game B two leaders have always selected Play-Red (plus one 39 times), and three followers choose Left and one Right in every single period.

5. Conclusions

This paper has provided experimental evidence on two different versions of the Dalek game where the forward induction equilibrium weakly iterately dominates the other Nash equilibrium in pure strategies. Our experiment aimed at gaining further insight on equilibrium selection dynamics and on the drift effect modeled by Binmore and Samuelson (1999).

First, our results show that the forward induction outcome is payoff-sensitive because it arises when we impose a change in the payoffs that does not modify the pure-strategy equilibria of the game and their being sub-game perfect. In the baseline treatment between the two games there is an increase in the forward induction outcome from about 30% to about 80% of plays. This result is confirmed with minor differences both in the risk-control

[15] Among the 10 "all-but-one players" of the baseline treatment, half of them played the "deviant" strategy in the first period usually obtaining a low pay-off since it was not an equilibrium strategy, whereas the remaining 5 played it rather randomly (33rd, 19th, 10th, 5th, and 12th period). This evidence reinforces the previous observation that in some cases initial strategies are reversed.

treatment and with an inversion in the sequence of the games. Still, our results support the hypothesis that the change in outcomes induced by the change in pay-offs is partly explained by their influence on the initial conditions of the learning process, that is, on the agents' initial beliefs. It is also noteworthy that a minority of subjects in the two sub-populations of players tends to hold to a fixed strategy, which *may or may not* be the modal strategy in the sub-population at the end of play, quite independently of changes in the frequency distribution of choices. Third, population learning plays a meaningful role, since there is a statistically significant and gradual change between the initial and final frequency distribution of choices in the two sub-populations. The frequency of dominated strategies tends to decrease.

These results provide some evidence *prima facie* consistent with the experimental relevance of the drift in the relation between learning and equilibrium selection within a randomly matched population. This outcome has been obtained by assuming that the first mover makes her decision first ("timing") but the second mover is not informed of the first mover's choice ("lack of observability"). Both players were informed of their position in the sequence and of the fact that the second player decided without knowing the decision of the first player. In our experimental setting this assumption may reinforce the focus of subjects' attention on the opponents' distribution of choices in previous interactions.

In our view this paper constitutes a first step in designing experiments aimed at capturing the role of learning in equilibrium selection and at controlling for the drift hypothesis. Future experiments could consider two groups of subjects that start playing a common modified version of the Dalek game that in subsequent repetitions gradually changes, as suggested in Binmore and Samuelson (1999). In this case players could not condition their initial behavior on different pay-offs and therefore cannot begin with systematically different initial play. Any difference in outcomes could be attributed to drift. A further experiment could attempt at testing the hypothesis that the initial selection of strategy distributions is explained by the theory of virtual observability. With reference to games otherwise identical to those considered in this paper, the experiment would consist in studying the effects of inverting the order of moves between sub-populations 1 and 2.

Table 1
Experimental Design Summary

Session	Treatment	No. of players	Run length (Game)
1	Baseline	12	40 (A) + 40 (B)
2	Baseline	12	40 (A) + 40 (B)
3	Baseline	12	40 (A) + 40 (B)
4	Baseline	12	40 (A) + 40 (B)
5	Baseline	10	40 (A) + 40 (B)
6	Risk control	12	40 (A) + 40 (B)
7	Risk control	12	40 (A) + 40 (B)
8	Reverse order	12	40 (B) + 40 (A)
9	Reverse order	12	40 (B) + 40 (A)

Table 2
Baseline Treatment Results (percentages)

		Game A		Game B	
		Initial	Average	Initial	Average
Player 1's choice	Exit	48.3	46.3	24.1	3.6
	Red	24.1	45.1	65.5	95.3
	Black	27.6	8.6	10.3	1.0
Player 2's choice	Left	51.7	53.5	79.3	86.5
	Right	48.3	46.5	20.7	13.5
Outcome	Exit, Left (A)	27.6	19.8	17.2	2.9
	Exit, Right (B)	20.7	26.4	6.9	0.7
	Play-Red, Left (C)	13.8	30.6	51.7	82.7
	Play-Red, Right (D)	10.3	14.5	13.8	12.7
	Play-Black, Left (E)	10.3	3.1	10.3	0.9
	Play-Black, Right (F)	17.2	5.6	0.0	0.2

Table 3
Risk Control Treatment Results (percentages)

		Game A		Game B	
		Initial	Average	Initial	Average
Player 1's choice	Exit	91.7	67.9	25.0	7.5
	Red	8.3	27.3	66.7	92.1
	Black	0.0	4.8	8.3	0.4
Player 2's choice	Left	25.0	49.4	41.7	92.1
	Right	75.0	50.6	58.3	7.9
Outcome	Exit, Left (A)	25.0	29.8	8.3	6.0
	Exit, Right (B)	66.7	38.1	16.7	1.5
	Play-Red, Left (C)	0.0	17.7	25.0	85.8
	Play-Red, Right (D)	8.3	9.6	41.7	6.3
	Play-Black, Left (E)	0.0	1.9	8.3	0.2
	Play-Black, Right (F)	0.0	2.9	0.0	0.2

Table 4
Reverse Order Treatment Results (percentages)

		Game A		Game B	
		Initial	Average	Initial	Average
Player 1's choice	Exit	33.3	50.0	25.0	21.3
	Red	53.8	46.0	66.7	74.0
	Black	8.3	4.0	8.3	4.8
Player 2's choice	Left	50.0	48.1	66.7	69.6
	Right	50.0	51.9	33.3	30.4
Outcome	Exit, Left (A)	16.7	20.0	16.7	13.8
	Exit, Right (B)	16.7	30.0	8.3	7.5
	Play-Red, Left (C)	33.3	26.5	50.0	53.8
	Play-Red, Right (D)	25.0	19.6	16.7	20.2
	Play-Black, Left (E)	0.0	1.7	0.0	2.1
	Play-Black, Right (F)	8.3	2.3	8.3	2.7

Table 5
Logit Estimation – Baseline Treatment

	Coefficient	S.E.	p-value
Game A			
Constant	-1.316	0.144	0.000
Lratio_{t-1}	2.786	0.250	0.000
	Log-likelihood -709.865	Baseline log-likelihood	-780.752
Game B			
Constant	1.056	0.701	0.127
Lratio_{t-1}	2.899	0.886	0.001
	Log-likelihood -151.227	Baseline log-likelihood	-156.096

N.B.: baseline log-likelihood is the log-likelihood of the estimation with only the constant as regressor.

Table 6

Learning – Baseline Treatment

	Coefficient	S.E.	p-value
Game A			
Constant	-0.194	0.083	0.021
Periodum	-0.090	0.118	0.445
	Log-likelihood -800.567	Baseline log-likelihood -800.860	
Game B			
Constant	2.909	0.187	0.000
Periodum	0.948	0.347	0.007
	Log-likelihood -176.478	Baseline log-likelihood -180.607	

N.B.: baseline log-likelihood is the log-likelihood of the estimation with only the constant as regressor.

FIGURE 1. ***Modified Dalek Game (Binmore-Samuelson 1999)***

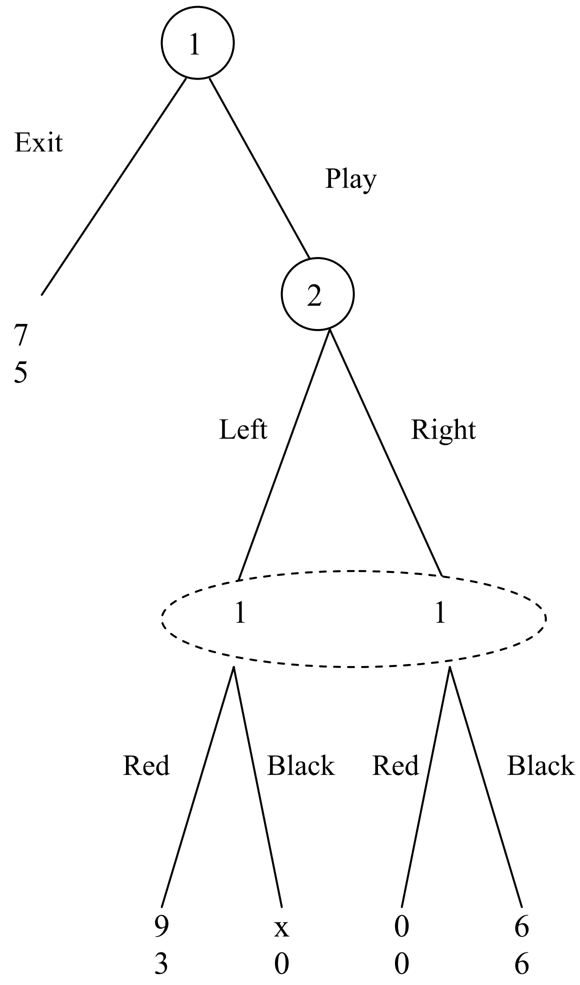


FIGURE 2a. ***Game A***

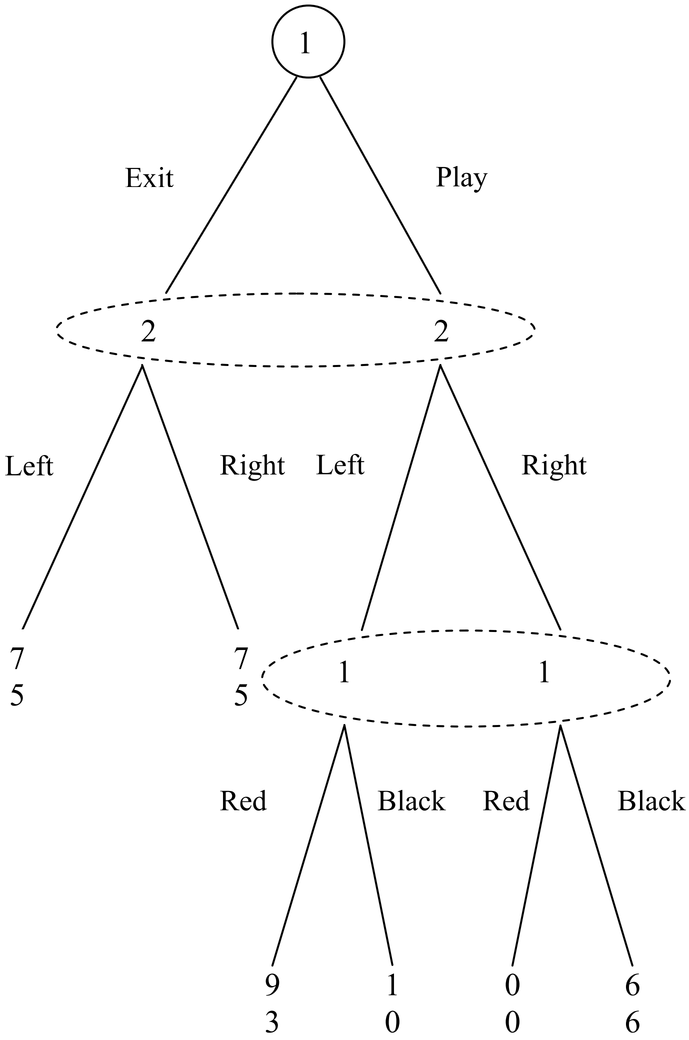


FIGURE 2b. ***Game B***

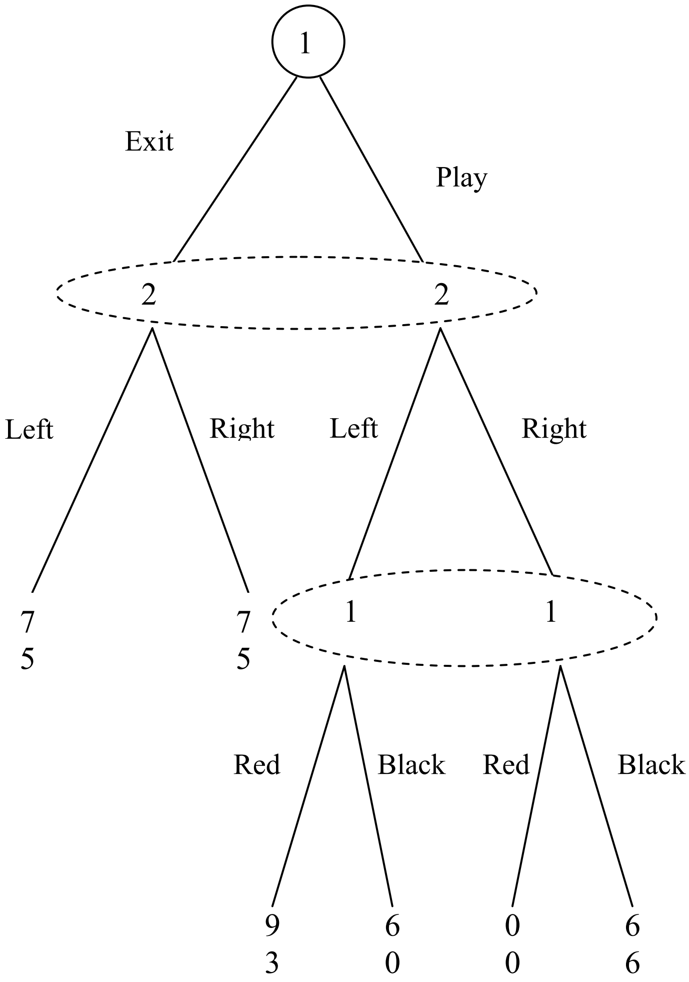


FIGURE 3. ***Games A and B in normal form***

Game A

Player 1 (blue)		Player 2 (yellow): Left	Player 2 (yellow): Right
	Exit	7, 5	7, 5
	Play – Red	9, 3	0, 0
	Play – Black	1, 0	6, 6

Game B

Player 1 (blue)		Player 2 (yellow): Left	Player 2 (yellow): Right
	Exit	7, 5	7, 5
	Play – Red	9, 3	0, 0
	Play – Black	6, 0	6, 6

Figure 4a. ***Choices of players 1 (Game A - baseline)***

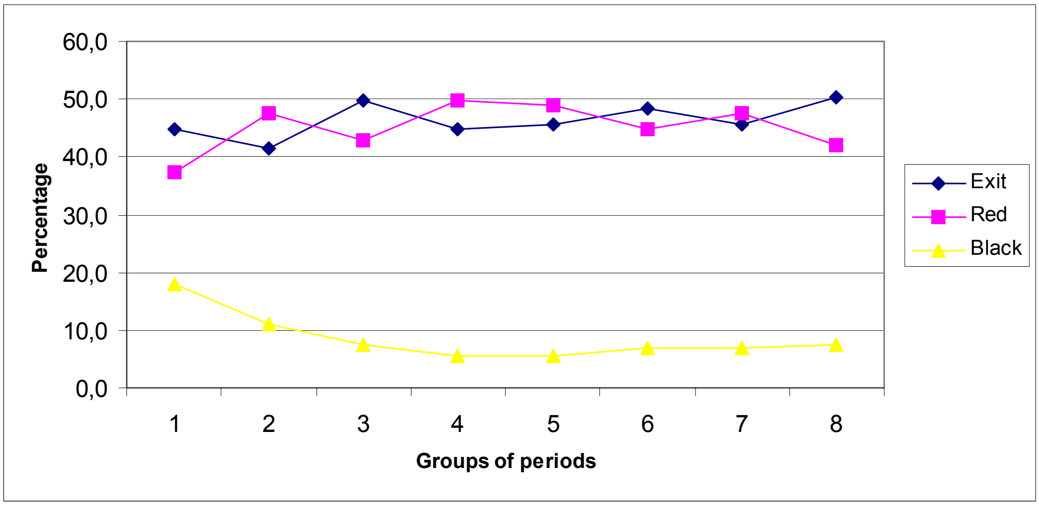


Figure 4b. ***Choices of players 2 (Game A - baseline)***

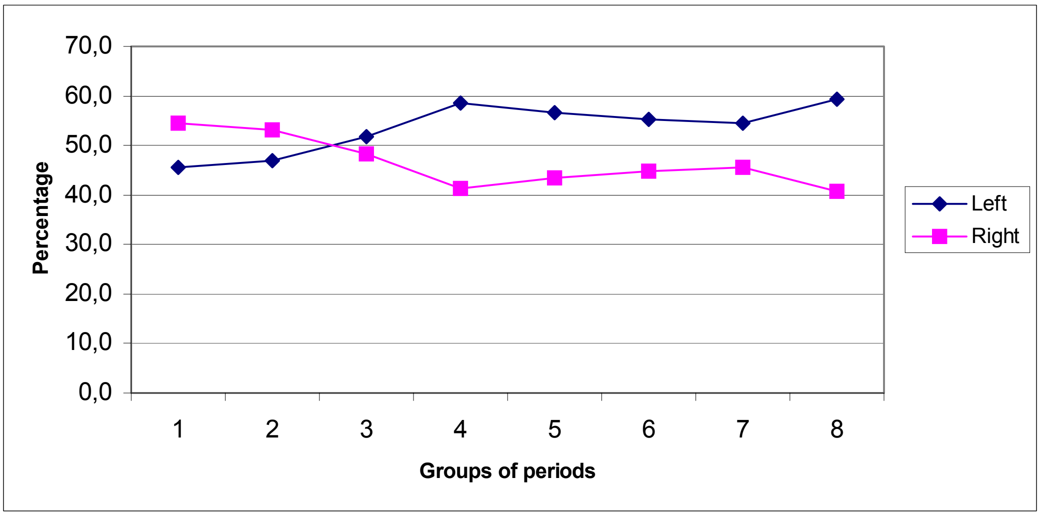


Figure 4c. ***Outcomes (Game A - baseline)***

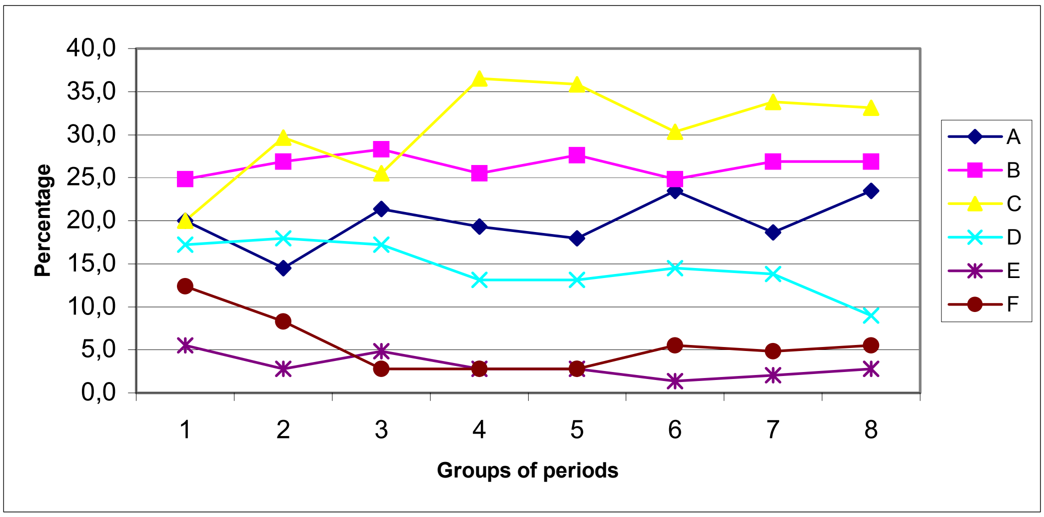


Figure 5a. ***Choices of players 1 (Game B -baseline)***

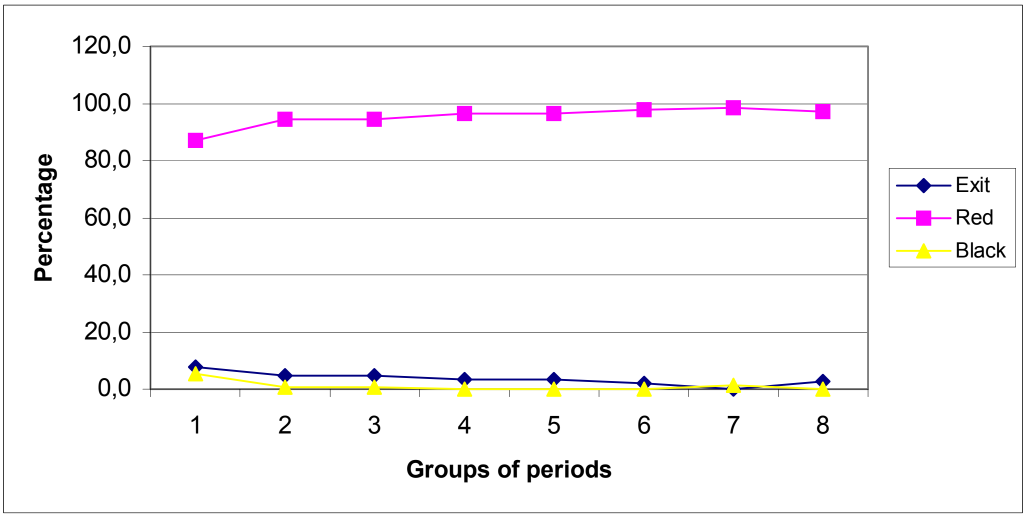


Figure 5b. ***Choices of players 2 (Game B - baseline)***

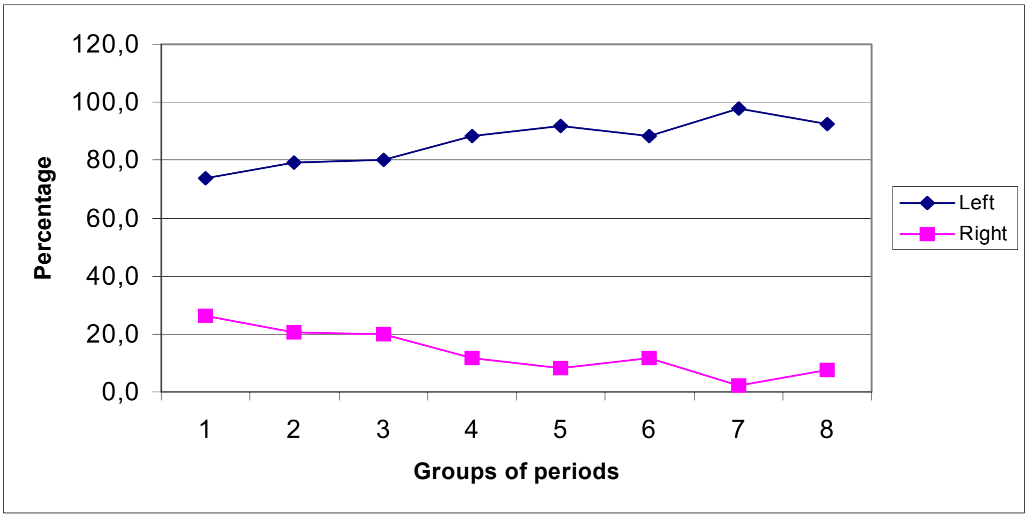


Figure 5c. ***Outcomes (Game B - baseline)***

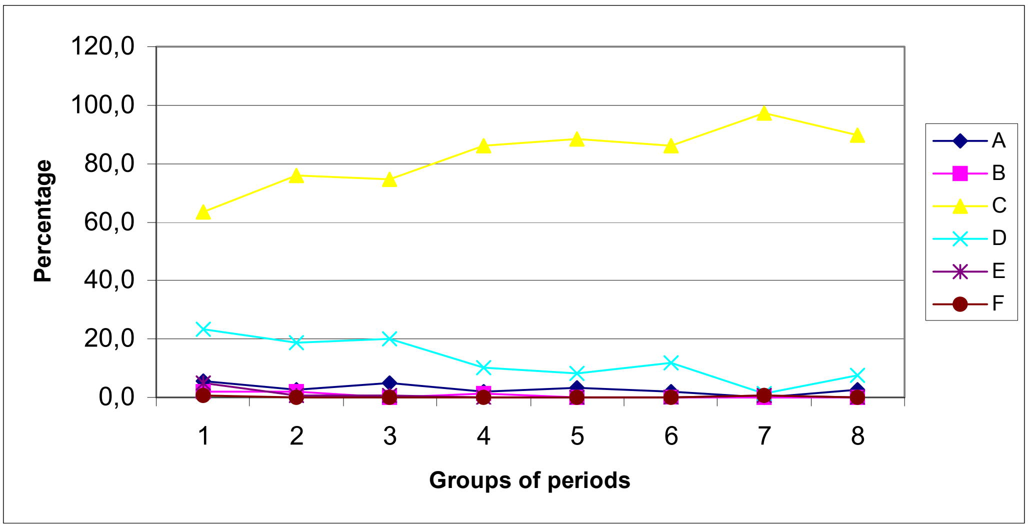


Figure 6a. ***Choices of players 1 (Game A – risk control)***

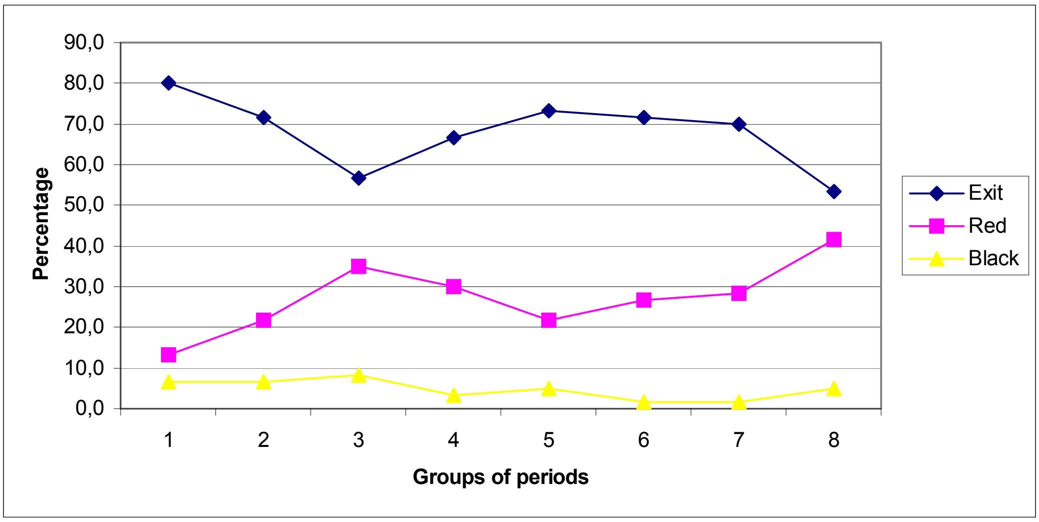


Figure 6b. ***Choices of players 2 (Game A – risk control)***

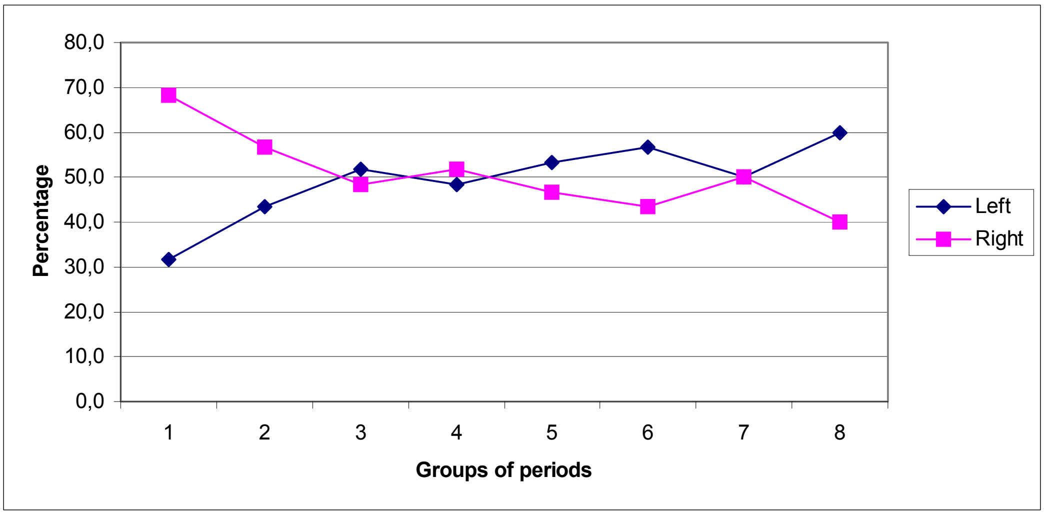


Figure 6c. ***Outcomes (Game A – risk control)***

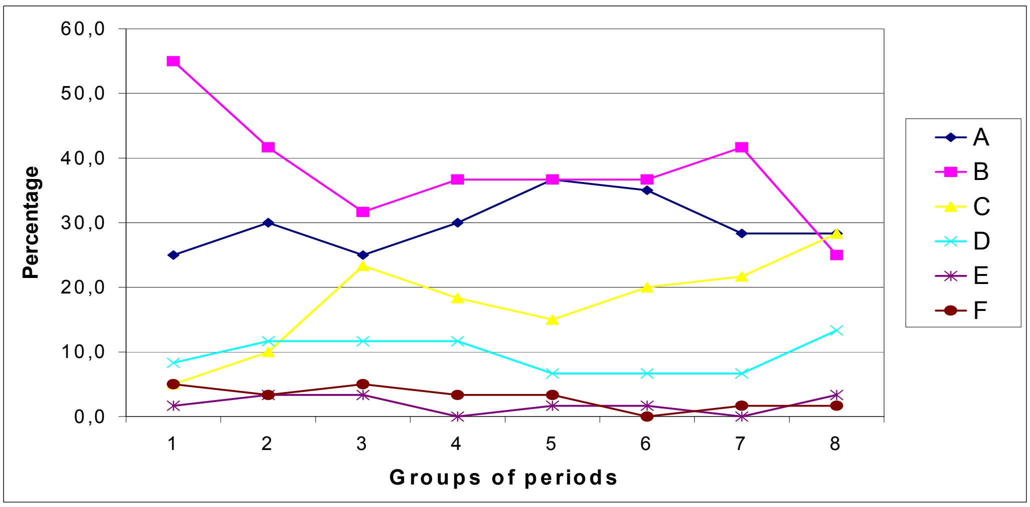


Figure 7a. ***Choices of players 1 (Game B – risk control)***

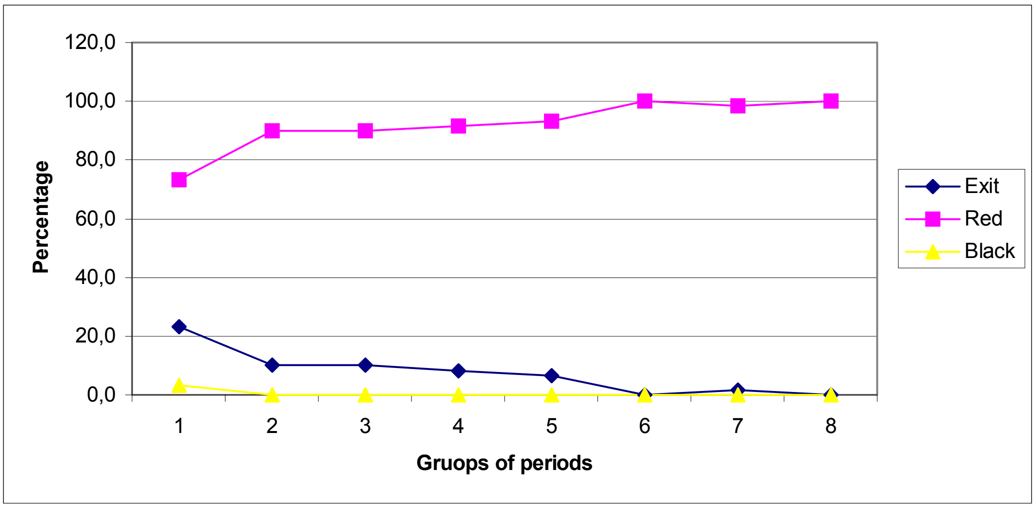


Figure 7b. ***Choices of players 2 (Game B – risk control)***

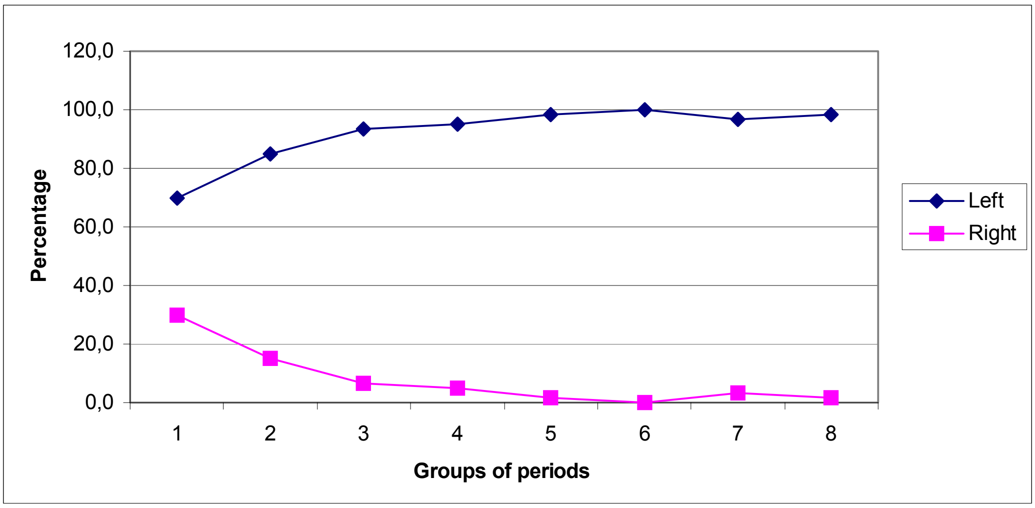


Figure 7c. ***Outcomes (Game B – risk control)***

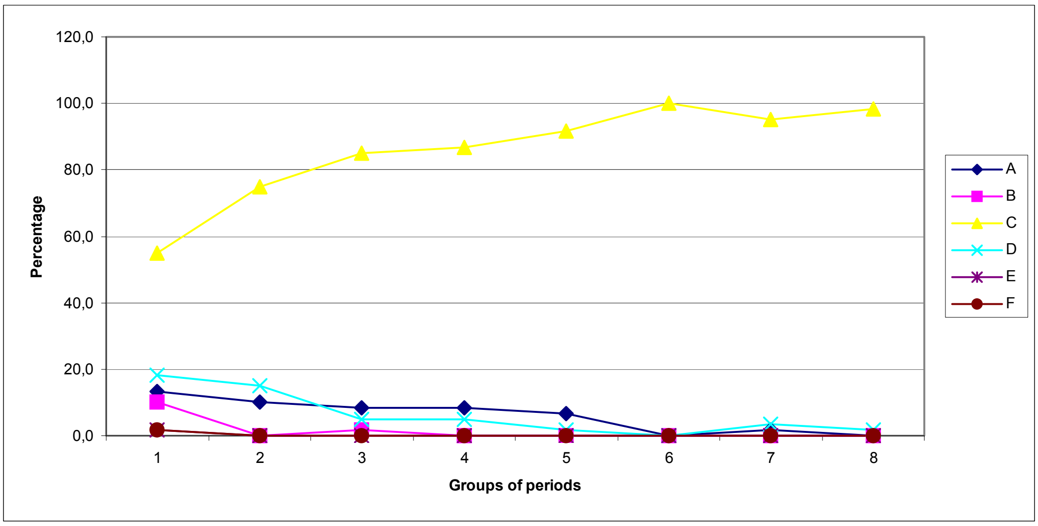


Figure 8a. ***Choices of players 1 (Game A – reverse order)***

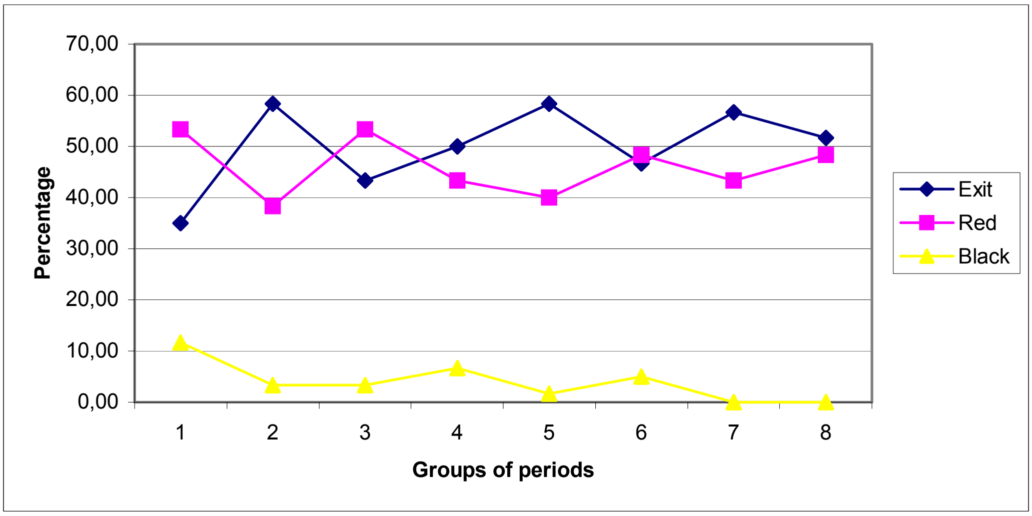


Figure 8b. ***Choices of players 2 (Game A – reverse order)***

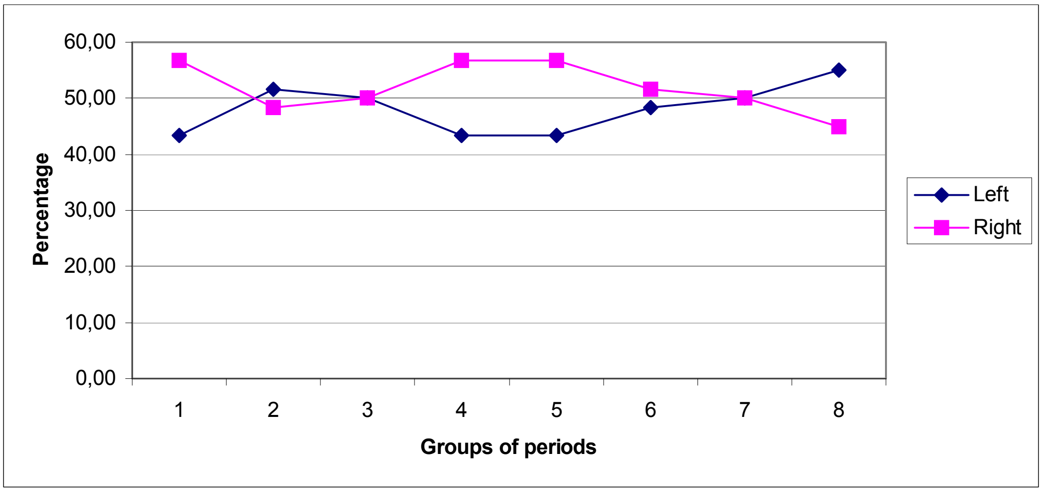


Figure 8c. ***Outcomes (Game A – reverse order)***

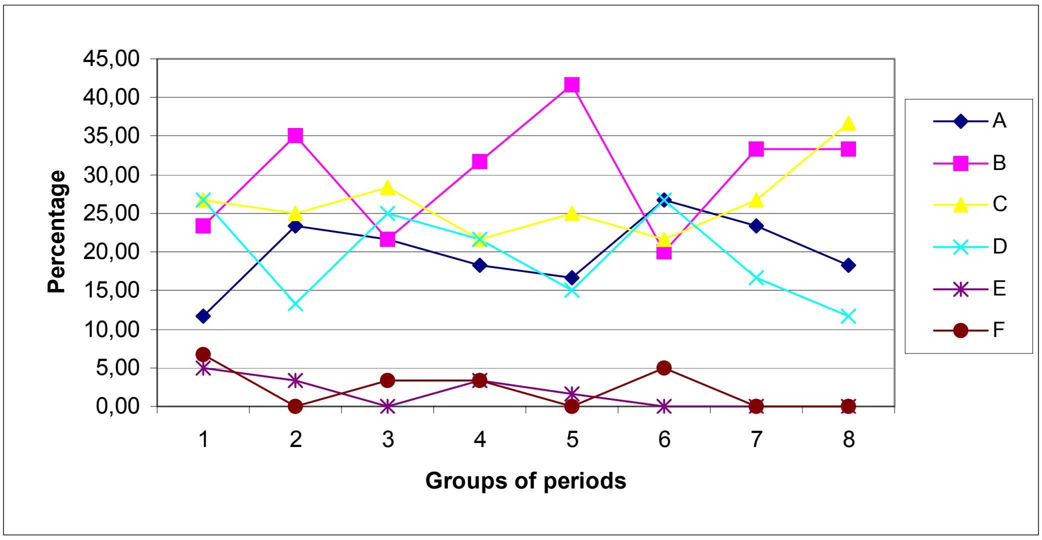


Figure 9a. *Choices of players 1 (Game B – reverse order)*

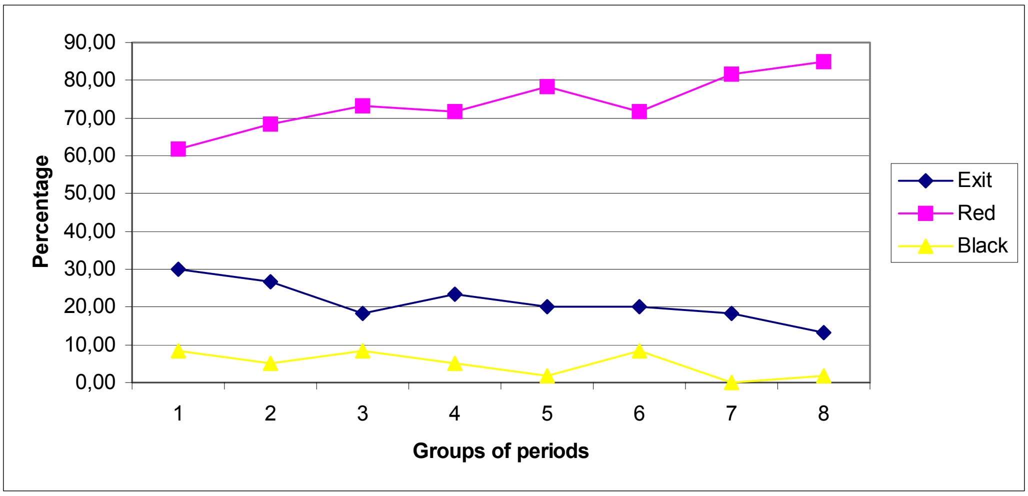


Figure 9b. *Choices of players 2 (Game B – reverse order)*

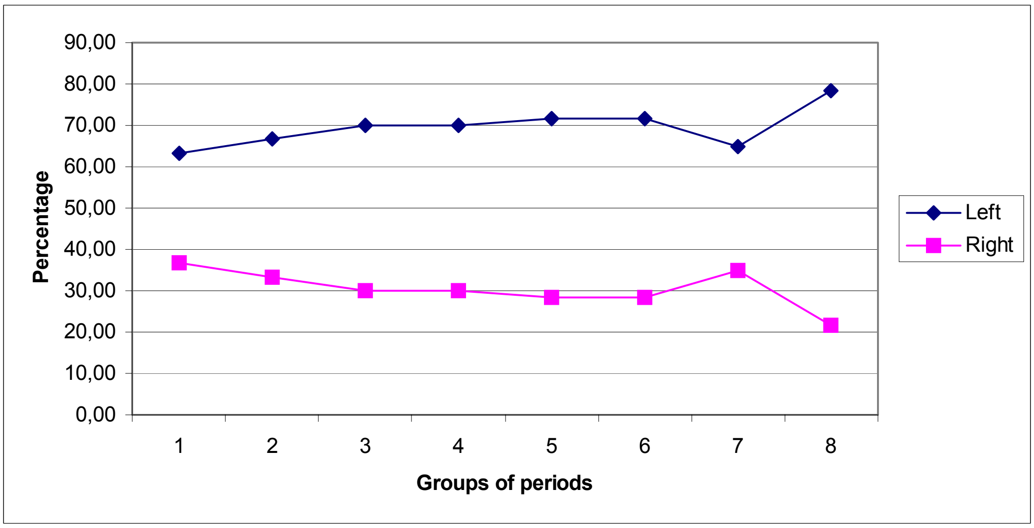
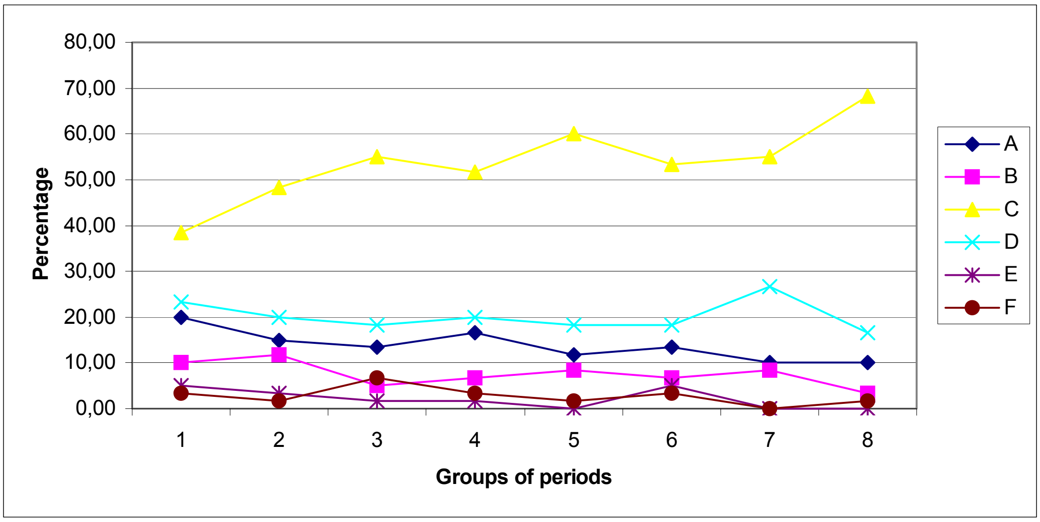


Figure 9c. *Outcomes (Game B – reverse order)*



APPENDIX

Translation of the instructions of the baseline experiment (distributed and read aloud)

This is an experiment on individual decision that was funded by the University of Siena. Making right choices you can earn an amount of money that will be given to you in cash at the end of the experiment. The experiment will last about one hour and during this period you should not talk to each other.

The experiment is made up by two sessions, respectively denoted as A and B. Both sessions consist of 40 repetitions of the game that we call periods. The computer will randomly divide in two groups, which we call blue group and yellow group, respectively. At the beginning of each period each blue player will be randomly paired with a yellow player. Pairings are drawn such that in each period each player would have the same probability of being matched with a player of the other group. In the screen points obtained by blue players are in blue, and those obtained by yellow players are in yellow.

At the beginning of each period a decision tree will appear on the screen (in the following sheet you can see the decision tree for session A). Blue players will decide whether to exit from the game - by clicking the button "exit" and getting the visualized points - or to play with the yellow player – by clicking on the button "play". In this case the blue player will choose between "red" and "black", without knowing the choice of the yellow player. In turn, the yellow player has to choose between "right" and "left", without knowing the choice of the blue player.

At the end of each period the computer will show the points obtained in that period, and how it depends on your and other choices. In addition, the choices of the members of the other group will be shown. Finally, results for the previous periods will be given.

Earnings will be the sum of a show up fee worth € 3 for the group blues and € 7 for the yellow group plus 2 eurocents times the number of total points you have earned during the experiment.

We start the experiment with a five-period trial. Points obtained during these periods will not be considered in calculating your final points and the associated individual payment.

Are there any questions?

References

Amershi, A.H., A. Sadanand, and V. Sadanand (1992). "Player Importance and Forward Induction", *Economics Letters*, 38, 291-297.

Balkenborg, D. (1994). "An Experiment on Forward- versus Backward-Induction", Discussion paper B268, SFB 303, University of Bonn.

Binmore, K. (1987–1988). "Modeling Rational Players", *Economics and Philosophy*, (Part I) 3, 179-214; (Part II) 4, 9-55.

Binmore, K., and L. Samuelson (1999). "Evolutionary Drift and Equilibrium Selection", *Review of Economic Studies*, 66, 363-393.

Binmore, K., J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked (2002). "A Backward Induction Experiment", *Journal of Economic Theory*, 104, 48-88.

Brandts, J. and C. A. Holt (1995). "Limitations of Dominance and Forward Induction: Experimental Evidence", *Economics Letters*, 49, 391-395.

Camerer, C.F., M. Knez, and R.A. Weber (1996). "Timing and Virtual Observability in Ultimatum Bargaining and 'Weak-Link' Coordination Games", Discussion Paper, California Institute of Technology.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1992). "Forward Induction in Coordination Games", *Economics Letters*, 40, 167-172.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1993). "Forward Induction in the Battle of the Sexes Games." *American Economic Review*, 83, 1303-1316.

Ellison, G. (1993). "Learning, Local Interaction, and Coordination", *Econometrica*, 61, 1047-1072.

Fishbacher, U. (1999). *Z-tree Manual*, Institute for Empirical Research in Economics, University of Zurich.

Friedman, D. (1996). "Equilibrium in Evolutionary Games: Some Experimental Results", *The Economic Journal*, 106, 1-25.

Guth, W., S. Huck, and A. Rapoport (1998). "The Limitations of the Positional Order Effect: Can It Support Silent Threats and Non-equilibrium Behavior?", *Journal of Economic Behavior and Organization*, 34, 313-325.

Harsanyi, J. C., and R. Selten (1988). *A General Theory of Equilibrium Selection in Games*, Cambridge Mass., MIT Press.

Kandori, M., G. Mailath, and R. Rob (1993). "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica*, 61, 29-56.

Kohlberg, E. and J.-F. Mertens (1986). "On the Strategic Stability of Equilibria." *Econometrica*, 54, 1003-1038.

Muller, A. and A. Sadanand (2003). "Order of Play, Forward Induction, and Presentation Effects in Two-Person Games", *Experimental Economics*, 6, 5-25.

Ponti, G. (2000). "Cycles of Learning in the Centipede Game", *Games and Economic Behavior,* 30, 115-141.

Rapoport, A. (1997). "Order of Play in Strategically Equivalent Games in Extensive Form", *International Journal of Game Theory*, 26, 113-136.

Roth, A. E. and I. Erev (1995). "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term", *Games and Economic Behavior*, 8, 164-212.

Roth, A. and M. Malauf (1979). "Game-theoretic Models and the Role of Bargaining", *Psychological Review*, 86, 574–594.

Uriarte, J. R. (2003). “A Model of Evolutionary Drift”, working paper IL 0103, Universidad del Pais Vasco.

Von Neumann, J. and O. Morgenstern (1947). *The Theory of Games and Economic Behavior* (2nd ed.). Princeton, Princeton University Press.

Weber, R. A. and C. F. Camerer (2001). “Timing and Virtual Observability in Ultimatum Bargaining and ‘Weak Link’ Coordination Games”, working paper, Carnegie Mellon University – California Institute of Technology.

Young, P. (1993). “The Evolution of Conventions”, *Econometrica*, 61, 57-84.