The Persistence of Discrimination: a View from the Lab

Antonio Filippin
(University of Milan and IZA)
.
University of Milan
Department of Economics
via Conservatorio, 7
20122 Milan, Italy
email: antonio.filippin@unimi.it

May 29, 2004

Abstract

This paper is an experimental analysis of the role played by workers' expectations in explaining the puzzling long-run persistence of observed discrimination against certain minorities in the labor market. The experiment provides evidence supporting the theoretical prediction that unequal outcomes may emerge due to disadvantaged workers' expectations of being discriminated against that survive even after a real effect, i.e. discriminatory tastes, has become negligible. A Self-Confirming Equilibrium is observed in which wrong expectations of being discriminated against affect subjects behavior and generate unequal outcomes that do not contradict their wrong beliefs. Even when learning takes place, disadvantaged subjects do not reach a situation of balanced outcomes. In other words, discrimination seems to have persistent effect and expectations are likely to increase the probability that observed discrimination lags behind its causes.

Keywords: Experiments, Discrimination, Workers' Expectations.

JEL classification codes: C92, J71, J15, D84.

The experiment described in this paper is (hopefully) an improvement of a previous version described in the paper "Discrimination and Workers' Expectations: Experimental Evidence." All the valuable suggestions I received on the previous version by Jose Apesteguia, Pierpaolo Battigalli, Thomas Bauer, Gary Charness, Andrea Ichino, Karim Sadrieh, and Karl Schlag as well seminar participants at the ESA 2003 and University of Siena have been very useful in helping me to design this new version and are therefore gratefully acknowledged also here. Needless to say, all the errors are mine.

1 Introduction

The goal of this paper is to provide experimental evidence concerning the role of workers' expectations as an explanation for the puzzling long-run persistence of observed discrimination against certain minorities in the labor market. The model used as a benchmark (see Filippin (2003a)) shows that *ex ante* identical groups of workers may be characterized by unequal outcomes in equilibrium due to their different beliefs. In particular, the model shows that unequal outcomes may arise when minority workers wrongly believe that they are discriminated against, even when employers do not do so either directly or statistically.

The underlying idea is not that beliefs differ because of sunspots. On the contrary, past experience of observed discrimination is likely to be the driving force. Empirical evidence supports that beliefs are actually different. The dataset used by Filippin and Ichino (2004) shows that, although men and women share very similar expectations about the magnitude of the gender wage gap, the importance they assign to the underlying causes differs. In fact, while a larger fraction of men think that "actual differences between men and women" matter, a larger fraction of women points towards the "employers' discriminatory tastes" as one of the causes for the expected gap.

However, a previous version of this experiment (see Filippin (2003b)) did not succeed in separating beliefs of the two populations. It provided some evidence supporting the theoretical predictions, but the different behavior of disadvantaged subjects vanished rather quickly during the treatment, failing to generate a Self-Confirming Equilibrium driven by wrong beliefs.[1] The experiment presented in this paper gets rid of the problem of separating the beliefs of the two populations. In particular, beliefs of the advantaged subjects are irrelevant as far as their optimal choice is concerned. In this way, only beliefs of the disadvanteaged workers become relevant.

The setting of the experiment replicates the model by means of an auction in which subjects are randomly assigned to two populations: "red" and "blue." In every lottery every participant has an endowment of 10 Euro cents and bids in order to win a prize worth 60 Euro cents. The prize is awarded to the highest bid(s). However, there are some "crazy computers" that never assign the prize to the Blues and this fraction is unknown to the subject. Does past experience works as a coordinating force? In other words, starting from a situation where there are many crazy computer, do the Blues still expect of not being awarded the prize when they do not know that the fraction of crazy computers is negligible? A positive answer to this question would provide evidence that historical factors are important in selecting one among different possible outcomes (path dependent equilibrium selection), pointing toward the existence of hysteresis. This would provide useful insights concerning the long-run persistence of discrimination in the labor market.

[1] This paper follows Davis and Holt (1993) and Roth (1994) using the following terms:
experiment: the collection of all data;
session: the collection of data involving the same group of subjects on the same day;
trial: a decision unit, one repetition of the game.

One session of the experiment perfectly replicates the Self-Confirming Equilibrium driven by wrong beliefs: after having learned rather quickly that the fraction of crazy computers at the beginning of the experiment is quite high (it is 87%, indeed) Blues start bidding 1 cent, which is the optimal choice given a sufficiently low chance of getting the prize. From that point onwards, both Blues always offer 1 cent and never receive evidence that contradicts the goodness of their choice, even though after the 10th trial the fraction of crazy computers is sharply and steadily decreased to 13%. Their lower bids make them less likely to win, leading to unequal outcomes that are consistent with wrong expectations that they were less likely to get the prize. In half of the sessions the Wilcoxon signed-rank test rejects that learning takes place, while the others are characterized by a significant although not perfect learning of the underlying percentage of crazy computers. What is important to stress is that, even when learning takes place, i.e. when players are not stacked in the Self-Confirming Equilibrium driven by wrong beliefs, the Blues do not reach a situation of balanced outcomes. In other words, discrimination seems to have persistent effect and expectations are likely to increase the probability that observed discrimination lags behind its causes.

The structure of the paper is the following. Section 2 describes the theoretical framework behind the experiment and summarizes the testable implications. Section 3 displays the design of the experiment as well as its procedure. Section 4 outlines the contributions to the literature that are related to this experiment. Section 5 contains the results, and Section 6 draws some conclusions.

2 Theoretical framework

This paper aims to provide experimental evidence concerning some testable implications derived from a model that analyzes the role of workers' expectations in explaining observed unequal outcomes in the labor market (see Filippin (2003a)). This section provides a summary of the model, emphasizing its testable implications. After the experiment is presented in section 3, several features are contrasted and compared in more detail with the corresponding parts of the model.

The model is formalized as a two-stage game of incomplete information in which populations of workers and employers are involved. In every constituent game, i.e. in every repetition of the game played by agents randomly drawn from their populations, one employer and two workers, one of whom is a minority worker, are randomly matched. The workers choose among three levels of effort (low, intermediate, high) and the employer promotes one and only one of the two workers after having observed their effort. The promotion is desirable because the job assigned to the promoted worker is assumed to be characterized by a lower cost of effort. Promotions also depend on employer's type, which captures the possible disutility of promoting a minority worker (discriminatory tastes).[2]

[2] Observable effort and incomplete information are the main features that distinguish this approach from the tournament literature started by Lazear and Rosen (1981). The two ap-

Workers know that there are two types of employer, but they do not know whether the employer they face is discriminatory or not. Also the distribution of types within the population of employers is unknown and workers have beliefs about it.

The importance of workers' expectations can be appreciated by comparing the equilibrium outcome in terms of promotions that may arise when minority workers overestimate the percentage of employers characterized by tastes for discrimination with a situation in which their beliefs are correct *ceteris paribus*. Assuming, for the sake of simplicity, that employers do not discriminate against minorities either directly or statistically, and that all the other sources of heterogeneity such as the distribution of ability among workers have been neutralized, unequal outcomes may still arise due to minority workers' wrong expectations. In other words, wrong beliefs about being discriminated against may be self-confirming. In this circumstance what happens is that in equilibrium minority groups, who expect being discriminated against, exert less effort on average, because of a lower expected return. This induces a lower percentage of promotions within minority workers, which in turn is consistent with their beliefs that employers are characterized by discriminatory tastes. On the other hand, when beliefs are correct, symmetric outcomes are observed.

It is worth stressing that a necessary condition for such a Self-Confirming Equilibrium is that beliefs of majority and minority workers differ. If both groups have wrong but similar beliefs about the fraction of discriminatory employers, their behavior will also be similar, and balanced outcomes in terms of promotions should be expected as long as there are no discriminatory employers. In a nutshell, this is what happened in the previous version of this experiment. The design of the experiment, in which a few treatments with a known discriminatory framework preceded a treatment an unknown fraction of discriminatory employers, did not succeed in separating beliefs of advantaged and disadvantaged subjects. Hence, when they played under (unknowingly) fair conditions, their behavior was similar and so did the outcomes. To get rid of the problem of separating beliefs in the laboratory, the new version of the experiment uses the trick of making the optimal action of the advantaged subjects independent of their beliefs about the presence of discriminatory employers. In particular the optimal action is the maximum level of effort, that corresponds to beliefs of no discrimination in the model. On the other hand, the return on choice of the disadvantaged subjects is kept tightly linked to their beliefs. The trick is simply to have two players from each population competing for the prize, instead of only one.[3]

Several implications arising from this model can be tested in the laboratory:

1. *When beliefs are similar, workers' behavior should also be similar.* Then, whether balanced or unbalanced outcomes will emerge depends on the actual fraction of discriminatory employers. In particular:

proaches share most of their predictions, one of which being that discrimination, when it is common knowledge, affects the two populations of workers in the same way (see also Section **??**).

[3] The implications of this trick are analyzed in details in Section 3.

a) when it is common knowledge that there is no discrimination, i.e. when the game is like a symmetric tournament, all workers should exert an inefficiently high level of effort;

b) when a known amount of discrimination affects workers' behavior, there should not be systematic differences across populations. In other words, the effort exerted by majority and minority workers should decrease in a similar way.

c) when the amount of discrimination is unknown and beliefs are similar the behavior should also be similar, regardless of beliefs being right or wrong.

All this predictions found strong support in the old version of the experiment. Therefore, the new version focuses on the following:

2. *Workers who overestimate discrimination exert a lower effort than workers characterized by correct expectations.* This is the key mechanism that might drive the labor market towards unequal outcomes even when discriminatory tastes have disappeared. In Filippin (2003a) a static framework is used and it is assumed that minority workers are those who might have wrong beliefs. Behind the static model there is an implicit dynamic: minority workers who have experienced direct discrimination for a long period continue to expect of being discriminated against even though discriminatory tastes have disappeared (hysteresis).

3 The Experiment

The experiment tries to capture the main features of the model that explores the role of workers' expectations in explaining observed unequal outcomes presented in the previous section. The game is much simpler than the model in order to be easily played. At the same time, the subjects are not made aware of the underlying economic relations being tested. Thus, keywords like discrimination, labor market, employer, worker, male and female are never used. This minimizes the risk that idiosyncrasies might enter the experiment and confound the results. Nevertheless, as reported in Section 5, gender keeps some relevance in explaining subjects behavior.

3.1 Sketch of the game

Participants are randomly divided in two populations: Red and Blue. Then, in every trial, groups of four subjects (two Reds and two Blues) are formed and play separate auctions. Every participant has an endowment of 10 Euro cents and decides how many cents to allocate as a sort of lottery ticket to get a prize worth 60 Euro cents. The minimum offer is one cent. Bets are not given back to the players, neither to the winner(s) nor to the loser(s). The prize is awarded to the highest bid and it is equally split if the highest bid is made by more than one player, unless the decision is taken by a "crazy computer," which instead assigns the prize to the higher bid(s) among the Reds, regardless of the bids made by the Blues. The actual fraction of crazy computers is unknown to the subjects, who are simply told that the fraction can range from 0% to 100% and can change in every trial.

In some sessions random matching between subjects takes places at the beginning of each trial. The players know that in every trial they face one opponent from their same population, and two opponents from the other population. Subjects are warned that it is possible to face the same opponents more than once during every session, but of course they do not know when. In other sessions, instead, the matching happens once at the beginning and lasts for the whole experiment.

3.2 Contrast and comparison with the model

Similarities with the model in Section 2 are straightforward only if one knows that this model is what the experiment aims to test. Colors (red and blue) are the equivalent of the observable characteristic (gender, race, etc.) that does not affect workers' productivity. The endowment of 10 cents is the same as the utility level when intermediate effort is exerted, i.e. the optimal level of effort when only the instantaneous utility in the first period is considered, as if promotions are not an issue. The amount bet plays the role of additional effort exerted to enhance the probability of being promoted. The prize stands for the promotion and, finally, the "crazy computers" play the role of discriminatory employers.

3.2.1 Populations and Number of Types

As already mentioned, red and blue labels are the equivalent of the payoff-irrelevant observable characteristic that distinguishes minority from majority workers. The color label is assigned randomly to every participant at the beginning and lasts for half of experiment. In the second part of the experiment color labels are switched, in order to prevent members of the blue population from feeling tempted to hinder the experiment.

The role of the population of employers is played by the computers, which implements the employers' equilibrium strategies in the model. The crazy computers never assign the prize to the members of the blue population. The "fair computers" instead assign the prize to the player who made the higher offer and they split the prize when bids are equal. Hence, only the blue players risk being discriminated against.

In the theoretical model it is necessary to assume that workers are of different types for the employers to have some uncertainty about their productivity in the second period. In the experiment the distinction of different types would make the game much more complicated, given that subjects are not familiar with the concept of payoff-type. A further appreciable gain in simplicity is that, since the computers directly play the equilibrium strategy of the employers, it is unnecessary to play the second stage of the theoretical model.

3.2.2 Utility function and Nash Equilibria

The utility function used in the model is not implemented directly in the experiment because it would be quite cumbersome to deal with it in the limited time-spell of the experiment (1 hour). However, the game sketched in Section 2 implies a simplified but very close version of it. In both cases players have the opportunity to give up some utility with certainty in exchange for an uncertain but higher return. In the model, supplying a high effort is a sub-optimal decision considering the instantaneous utility of the first period only, but such a loss of utility can be more than counterbalanced if the worker is promoted, since the job assigned to the promoted workers is characterized by a lower cost of effort. On the other hand, if the worker is not promoted (s)he suffers a net loss of utility with respect to the case in which (s)he had chosen the "safe" option, i.e. intermediate effort. The risk of not being promoted has its counterpart in the possibility of bidding without getting the prize in the experiment. However, also in this case the player has a "safe" option, corresponding to low effort in the model, which is bidding one cent.

It is useful to starting the analysis of the Nash Equilibria in the experiment, where the fraction of crazy computers is unknown, from two benchmark situations: two auctions where it is known that the prize is assigned by an unbiased or by a crazy computer, respectively.

In the first case, the utility function of all the four players is the same

$$u_i(s_i) = \begin{cases} 10 - s_i & \text{if} \quad s_i < \max(s_{-i}) \\ 10 - s_i + \frac{60}{n} & \text{if} \quad s_i \geq \max(s_{-i}) \end{cases} \tag{1}$$

where s_i is the strategy, i.e. the bid, of a generic player i, $-i$ stands for the opponents and $n \leq 4$ is the number of winners. Clearly,

$$10 - s_i + \frac{60}{n} > 10 - s_i$$

holds whatever the outcome of the auction is in terms of number of winners $n \leq 4$. This means that every player has the incentive to overbid by one cent, whenever possible, the best offer of her opponents in order to win the entire prize

$$BR_i(s_{-i}) = \{s_i : s_i = \min[10; \max(s_{-i}) + 1]\}.$$

However, since this holds for all players, the only Nash Equilibrium of the game is $BR_i(s_{-i}) = \{10\}$. This Nash equilibrium is not efficient, because all players would be better off if all would bid one cent, but in that case each player would have the incentive to deviate bidding a slightly higher amount. This is similar to what happens in the theoretical model for the combination of parameters that displays a Bayes-Nash Equilibrium with both workers supplying a high effort. Given that the model imposes that one and only one worker is promoted, if both workers offered an intermediate effort, the probability of being promoted would be the same and both would have a higher utility in the first period. However, neither worker would be maximizing his/her utility because supplying a high effort would be a profitable deviation.

How does the picture change when the decision rule of a crazy computer who always assigns the prize among the Reds is considered? The situation is dramatically different for the Blues, identified by the superscript B, whose utility function simply becomes

$$u_i^B(s_i^B) = 10 - s_i^B \qquad \forall s_{-i},$$

trivially leading to $BR_i^B(s_{-i}) = \{1\}$. Bidding one cent is a dominant strategy for both blue players, just as intermediate effort is a dominant strategy in the model for the minority worker when there are discriminatory employers only. The behavior of the Reds is instead unaffected by the type of computer assigning the prize. Looking at the utility function of the Reds, identified by the superscript R,

$$u_i^R(s_i^R) = \begin{cases} 10 - s_i^R & \text{if } \ s_i^R < \max(s_{-i}^R) \\ 10 - s_i^R + \frac{60}{n} & \text{if } \ s_i^R \geq \max(s_{-i}^R) \end{cases} \tag{2}$$

it holds the same argument made when the computer was unbiased. Even though the Blues cannot be awarded the prize (hence $n \leq 2$ in this case), the Reds still compete with each other, and they still have the incentive to overbid the best offer of the opponent

$$BR_i^R(s_{-i}) = \left\{ s_i^R : s_i^R = \min\left[10; \max(s_{-i}^R) + 1\right] \right\}.$$

However, since this holds for both Red players, it turns out that $BR_i^R(s_{-i}) = \{10\}$, like in the case of an unbiased computer. Therefore, the only Nash Equilibrium of the auctions ruled by a crazy computers happens when Reds bid 10 and Blues bid 1.

Having analyzed the two simple benchmark situations, it is now easier to characterize the more general framework of the game, i.e. when the fraction of crazy computers is unknown and players have beliefs about it. The behavior of the Reds is not affected by the type of computer. Therefore, the fraction of crazy computers, and consequently their beliefs about it when such a fraction is unknown, do not play any role in shaping their behavior. This is trick that differentiates this version of the experiment from the previous one. Since separating beliefs of Reds and Blues in the lab turns out to be very difficult, including two players from each population instead of only one induces the Reds to play always 10, regardless of their beliefs about the fraction of crazy computers. What is left to analyze is the behavior of the Blues when the fraction of crazy computers is unknown. Given that Reds plays 10 in any case, their utility becomes

$$u_i^B(s_i^B | s_{-i}^R = 10) = \begin{cases} 10 - s_i^B & \text{if } \ s_i^B < 10 \\ \left(10 - s_i^B\right) \mu_i^B + \left(10 - s_i^B + \frac{60}{n}\right)\left(1 - \mu_i^B\right) & \text{if } \ s_i^B = 10 \end{cases}$$

where μ_i^B represent player i's expected fraction of crazy computers. As long as a Blues offer less than 10, she is sure, given the opponents' strategies, that she will lose the prize. The only strategy lower than 10 that makes sense is of course the safe option, i.e. bidding one cent, which gives a utility equal to 9. Therefore, solving

$$\frac{60}{n}\left(1-\mu_i^B\right)\geq 9$$

for μ_i^B allows to find the threshold level of the expected fraction of crazy computers consistent with bidding 10 as a best reply. The result is

$$\mu_i^B \leq \frac{60-9n}{60},$$

with $n=\{3;4\}$. We are computing the expected earnings of a Blue player using her subjective probability of facing an unbiased PC, which weights her payoff in that circumstance. Hence also this Blue player must be among the winners, besides the two Reds. n is therefore at least equal to three in the inequality above. n can also (but does not need to) be equal to four, according to the choice of the other Blue player of whether bidding 10 or not. The resulting threshold are $\mu^B \leq 0.4$ with $n=4$ and $\mu^B \leq 0.55$ and with $n=3$. This means that whenever a Blue believes that the fraction of crazy computers is lower than 40%, bidding 10 is her best reply to the Reds bidding 10, regardless of the choice of the other Blue. On the other hand, when the expected fraction of crazy computers is higher than 55%, the optimal choice of a Blue is to bid the minimum, one cent, regardless of the choice of the other Blue. For values of beliefs between the two threshold levels, the optimal choice depend also on the strategy of the other Blue. Summarizing, when a Blue think that there are sufficiently many crazy computers, the optimal choice is the safe option, i.e. bidding one cent. Therefore, differently from what happens to the Reds, beliefs of the Blues are crucial in shaping their behavior.

3.3 Design and Procedure of the Experiment

One aspect of the experimental procedure needs to be stressed. All the treatments were proposed within each of the seven sessions of the experiment. Hence, all the subjects played facing the whole set of parameters. This procedure implies potential carryover effects from one parameter set to the others, as well as confounding factors arising because of framing, learning and fatigue. However, testing the existence of carryover effects (hysteresis) is one of the primary goals of the paper, and therefore such an approach is necessary. Moreover, an econometric approach to the analysis of the data allows us to control for any observable and/or unobservable individual characteristic that might affect the choices of the participants during the experiment, including framing, learning and fatigue. To minimize the role of confounding factors, simultaneous parameter changes are avoided.

What follows is a sketch of the rules and the procedure of the experiment, which has been run using the zTree software.[4] We recruited subjects from under-

[4] The zTree software was developed at the University of Zurich, Institute for the Empirical Research in Economics (see Fischbacher, 2002).

graduate courses at the University of Milan. Most of the subjects were inexperienced. Participants were first randomly assigned numbers and seats. Subjects were told that their physical identity was not associated with their choices during the experiment, the subjects' numbers being their personal identification. They were given written instructions, the first part of which, concerning the matching procedure and the rules of the game without crazy computers, were also read aloud by the experimenters, who stressed that earnings are a function of subjects' decisions.[5]

Quiz 1. After questions are raised by subjects, a quiz is run to test their comprehension of the game without crazy computers. First, participants are asked to compute the earnings of a fictitious player, starting from her bid and the bids of her opponents that were displayed on the screen. Then, given three different sets of bids, they are asked to answer which player(s) is (are) awarded the prize. If wrong answers are given, the subsequent screen shows the subject the correct answer and, in the case of earnings, the way to compute them. Subjects are invited again to ask questions about anything unclear.

Introduction of crazy computers. After Quiz 1, the participants are told that there are also some "crazy computers," i.e. computers that assign the prize to the highest bid(s) within the red population and never to the Blues, regardless of their bids. They might face either unbiased or crazy computers, but obviously they are not told at the beginning of each trial whether the computer running that trial is crazy or not. Moreover, they are not told even the fraction of crazy computer during the experiment, but it is stressed that there is a random matching between groups of four subjects and computers in every trial. The reason is to make clear that it is not the computer in front of which they are sitting that may be crazy, and that the type of computer they face in one trial does not convey additional information besides that on the overall fraction of crazy computers.

Participants are warned that the maximum attention has been paid in order that every subject might get the same expected reward, with differences depending on participants' actions only. Accordingly, each player belongs to the blue population for half of the trials and to the red population for the other half. The reason is to prevent members of the blue population from feeling tempted to hinder the experiment.

Quiz 2. Subjects are then asked to answer another short quiz to test their comprehension of the game when crazy computers are introduced. Given three sets of bids, divided by color, subjects are asked to identify the player(s) to whom the prize is awarded according to the type of computer. If wrong answers are given, the subsequent screen shows the subject the correct answers. Subjects are invited again to ask questions about anything unclear.

Assignment to red or blue population. The color of the population is then randomly assigned to every participant by means of an algorithm, in such

[5] Both the matching procedure, which varies across sessions, and the rules of the game are explained in more detail below.

In the Appendix, the instruction of one sessions are reported as an example. The instructions of the other sessions are sligthly different, according to the different setting.

a way that unobserved and uncontrolled characteristics are not correlated with the focus variables. The color assigned lasts for the first half of the experiment, then it is switched in the second half.

The game. After the quizzes the game starts. In every repetition of the game, each subject is endowed with 10 cents. Within each group of four, 2 Reds and 2 Blues, each player decides how much to bid in order to win a prize of 60 cents. If the prize is assigned by an unbiased computer, the highest bid wins the prize, and if the highest bid is made by more than one player the prize is equally split. If, on the other hand, the prize is assigned by a crazy computer, the choice of the winner is limited within the Reds, following the same rules. In both cases, bids are given back neither to the winner(s) nor to the loser(s).

After each participant has decided the computer displays to each player:

a) how much the four opponents bid, divided by color, in the auction in which she is directly involved;

b) whether she wins, shares or loses the prize;

c) (only in session 6 and 7) how many players win the prize in the auction in which she is directly involved.[6]

The fraction of crazy computers is unknown to the players. They are just told that the fraction of crazy computers can range between 0% and 100% and that it may change from trial to trial. This makes the theoretical predictions uncertain, since the best replies of the Blues depend on their beliefs, while Reds are expected to bid always 10 cents, as shown in the previous subsection. The actual fraction of crazy computer is equal to 87% during the first ten trials, and then it is sharply decreased to 13% for the other trials in the first half of the treatment. If this fractions were known, the behavior that could be expected is that the Blues bid 1 cent in the first ten trials, and 10 cents afterwards.

The Self-Confirming Equilibrium driven by wrong minority workers' beliefs corresponds in the game to a situation in which, after the fraction of crazy computers is decreased to 13%:

- Reds bid 10 cents;
- Blues, after having experienced discrimination in the first 10 trials, still think that there are sufficiently many crazy computers (at least 55% as explained above). Given this beliefs bidding one cent is a best reply. However, if Blues bid 1 cent, also unbiased computers will not assign them the prize.
- The prize is awarded to the Reds, and this is consistent with Blues wrong beliefs that there are many crazy computers.

[6] The role of this piece of information is to make it easier to learn the type of the computer for those who loses the prize. An example is useful to make things clear: suppose that the two Reds and one of the Blues bid 10, while the other Blue bid 1. With only the information in a) and b) the Blue that bid 1 cannot retrieve the type of the computer, while this is possible with the additional information in c). In fact, knowing that three players shared the prize she can infer that also the other Blue was assigned the prize, and therefore the computer must be unbiased.

Table 1:
Summary of the main features of the 7 sessions

Session	1	2	3	4	5	6	7
Nr. of participants	12	12	4	12	4	12	8
Random matching of players	yes	yes	no	yes	no	yes	yes
Nr. of trials with the same label	25	25	25	30	30	30	30
Learning even when bidding low	no	no	no	no	no	yes	yes

After the first half of the experiments, before color is switched, subjects are asked to report their beliefs about the actual fraction of crazy computers, by choosing between five equally sized intervals (0-20%; 20-40%;...) for each group of five trials (1-5; 6-10; ...). Expectations are elicited implementing a lottery in which each subject has a probability of winning that is proportional to the number of times in which her beliefs are correct.

Finally, the color is switched and the game is repeated identically.

The structure of the game is the same in all the seven sessions, which however differ along three possible dimensions:

a) the length of the experiment: in session 1, 2, 3 there are 25 trials before the color is switched, in the other sessions instead 30 trials;

b) in all the sessions apart from session 3 and 5 the groups of four players are formed with random matching before every trial. Sessions 3 and 5 instead are formed by four participants only, and therefore the situation is equivalent to a matching made once and for all at the beginning of the experiment;

c) as mentioned above, in session 6 and 7 players are given additional information after each trial, namely the number of winners. This makes possible for players to learn the type of the computer even thought they make a low offer, provided that one of the Blues matches the highest offer. The main features that characterize each session are summarized in Table 1.

Questionnaire. At the end of the experiment a questionnaire is proposed, reminding participants that their physical identity was not associated with their choices and their answers during the experiment. Questions concerned academic as well as personal information. In section 3.4 some descriptive statistics of the pool of subjects are summarized.

64 subjects participated in the experiment. The sessions lasted 60 minutes and were composed of a minimum of 4 and a maximum of 12 subjects. Euro cents were the currency used during the experiment. Earnings ranged between 8.5 and 15 Euros (11.1 on average).

3.4 Sample description

From the information collected by means of the final questionnaire, it turns out that females are over-represented in our sample (68% vs. 32%), and that the average age of the pool is about 22 years. Most of the participants (81%)

comes from the School of Political Sciences, and is enrolled in the third year of the degree program. The final mark at the exit of secondary school was chosen as a proxy for a student's ability; the variable has been re-scaled in the range [0,1]. More than two thirds of the sample come from high schools (*licei*) and one fourth from technical schools (*istituti tecnici*). Two specific questions were asked concerning political and religious orientation. An ordered scale from 0 to 5 has been used to ask subjects about their political orientation (0=left; 5=right), without any label on each possible choice. Roughly 70% of the subjects report themselves as being center-left, i.e. they chose a value from 0 to 2, and 30% center-right. The average choice is 1.98 while the median choice is 2. With respect to religion, the subjects have been asked to choose from three alternatives: "believer and churchgoer," "believer but not churchgoer," "non-believer." The proportion of the last occurrence was slightly more than one third.

4 Literature Review

Although the role of workers' expectations in explaining unequal outcomes has never been the focus of experiments, several contributions to the literature are relevant as far as this experiment is concerned. They can be divided into three groups:

1. Discrimination and asymmetric tournaments.
2. Sunspot and hysteresis.
3. All-pay auctions.

4.1 Experimental Studies of Discrimination and Asymmetric Tournaments

Experiments closely related to the experiment presented in this paper are those concerning either statistical discrimination or asymmetric tournaments. This subsection concentrates on experiments based on economic factors; a survey of many other experiments based on group identification or status can be found in Anderson, Fryer and Holt (2002).

The literature concerning experimental studies of discrimination is thin and in general not directly related to the experiment presented in this paper, with a few exceptions. Fryer, Goeree and Holt (2002) describe the results of experiments that may produce (and sometimes do) a pattern of experience-based discrimination consistent with the statistical discrimination models proposed by Arrow (1973) and Phelps (1972). Employers have to decide whether to hire or not workers from two otherwise identical populations, "green" and "purple". The hiring decision is affected by an observable test score, which in turn depends on a worker's (unobserved) investment decision, like education or training. The cost of investing is random and it is set to be systematically higher for the workers of one population during the first ten out of sixty rounds, while from the eleventh onward it is drawn from the same distribution. Moreover, players have access to aggregate information, given that the average investment and hiring percentages for the workers of each color are displayed at the end of each round. The authors find that a different average investment emerges, and then a lasting and self-reinforcing mechanism operates in such a way that fewer workers of that group are hired, the fraction of that group of workers investing decreases even further and so on, leading to multiple equilibria with discrimination. There are certain dimensions in which this experiment should be explicitly compared with that presented below. In particular, it is worth noting that, similarly to the experiment described in this paper:

a) There is a real effect (the different distributions from which investment costs are drawn) that is withdrawn during the experiment, but that have long-lasting effects (hysteresis);

b) There is an endogenous decision (investing or not) that makes *ex ante* equal populations potentially different in equilibrium.

Hargreaves-Heap and Varoufakis (2002) use the Hawk-Dove game to show that starting from two populations that differ because of a payoff-irrelevant observable characteristic only (red and blue label), different roles associated with different payoffs (i.e. discriminatory conventions) may emerge. Löhm (2000) finds in a Battle of the Sexes experiment that females are more likely to be discriminated against by other females.[7]

The literature concerning experimental studies of asymmetric tournaments is more established and some papers can fruitfully be used as a benchmark, in particular Schotter and Weigelt (1992) and Bull, Schotter and Weigelt (1987). The former presents an experiment aiming to test the theoretical predictions of the asymmetric tournament theory as presented by O'Keeffe, Viscusi, and Zeckhauser (1984). In particular, they focus on the predicted trade-off between equity and efficiency associated with affirmative actions, finding contrary evidence.

From the theoretical point of view, asymmetric tournaments have many things in common with the model presented in Section 2. In line with the old saying that different opinions are necessary for a horse race to take place, both involve uncertainty. What distinguishes them is the fact that effort is not perfectly observable in the asymmetric tournament literature, while incomplete information about the opponents' type-strategy set characterizes in Filippin (2003a). Furthermore, the two approaches share most of their predictions, in particular that the behavior of advantaged and disadvantaged workers should change in a similar way when discrimination is common knowledge.

The experiment presented by Schotter and Weigelt (1992) can be used as a benchmark also from the methodological point of view. Two points are partic-

[7] Other experiments concerning statistical discrimination have been proposed by Davis (1987) and Anderson and Haupert (1999), both relying on exogenous differences that characterize the two populations. The former finds weak evidence that the larger population has better outcomes. The latter provides evidence that workers belonging to a population characterized by a lower average innate productivity are less likely to be hired, with the likelihood depending on the cost of discovering the individual type. Strictly speaking, it can be argued that the framework of these experiments cannot be classified as discrimination.

ularly relevant:

a) The authors want to avoid carryover effects from one treatment to another. Consequently, each subject is allowed to participate in one treatment only. In the present paper, instead, the main goal is to test the existence of hysteresis, and therefore carryover effects are part of the picture. Hence, the treatments are designed in such a way that every subject faces both a symmetric game and situations characterized by discrimination;

b) In the experiment proposed by Schotter and Weigelt (1992) players are matched once and for all within every session. This is more likely to lead to cooperation, to foster strategic interaction, or at least to lower the chances to learn. Such effects are instead more likely to disappear with a random matching repeated before every period (see Duffy and Ochs (2003)). As explained in Section 4 in more details, this experiment involves both sessions with random matching and sessions with matching once and for all.

The experiment just described closely follows an earlier experiment by Bull, Schotter and Weigelt (1987), where asymmetries and affirmative actions were not the main focus. It is worth noting that both experiments report a tendency of disadvantaged workers to over-supply effort in uneven tournaments, as if asymmetries elicit greater effort.[8]

4.2 All-pay auctions

All-pay auctions are characterized by the fact that bids are given back neither to the winner nor to the loser. The model behind the experiment is related to an all-pay auction, insofar as there is no compensation for the loss of utility that a non-promoted worker suffers when he exerts an effort higher than the optimal level considering only the instantaneous utility function in the first period, i.e. if promotion is not an issue. In an all-pay auction the prize goes to the highest bidder, so that each player has the incentive to overbid the others, as long as this ensures a positive payoff.[9] When the value of the prize exceeds the sum of the endowments, like in this experiment when the fraction of crazy computers is sufficiently low, an equilibrium in pure strategies exists, implying full dissipation of the endowments. Otherwise, in symmetric all-pay auctions, the result that the sum of the expected bids equals the value of the prize is supported by mixed strategies equilibria (Baye, Kovenock and de Vries, 1996). Rational agents never over-dissipate the value of the rent if they have the opportunity to bid zero. However, a relaxation of the rationality via the possibility of decision errors is enough to support a theoretical framework where over-dissipation can be observed (Anderson, Goeree and Holt, 1998) consistently with experimental evidence like Davis and Reilly (1998).[10]

[8] Asymmetric contests are defined "uneven" when agents are different, and "unfair" when contestants are identical but the rules favor one of them.

[9] The literature on symmetric rank-order tournaments started by Lazear and Rosen (1981) shares some of the features concerning all-pay auctions.

[10] When within such a framework subsequent bids are allowed before the prize is assigned, it is easy to observe that bidding spirals out of control, as in the Dollar Auction Game presented

The promotion game with discrimination in Filippin (2003a), tested in this experiment, can be classified as an all-pay auction, since bids are sunk. However, the presence of discriminatory employers prevents the prediction of full dissipation of the endowment from applying straightforwardly, as seen in section 3.

4.3 Sunspots and Hysteresis

The theoretical sunspot model postulates that agents believe that a variable, which is in fact unrelated to the economy, has real effects, and shows that such beliefs can induce the agents to behave in a manner that provides support for the postulated beliefs. Sunspots were introduced in the laboratory by Woodford (1990), who shows that cyclic sunspot equilibria can asymptotically emerge in an OLG framework when agents follow some adaptive learning schemes. Marimon, Spear and Sunder (1993) do not find evidence that sunspot equilibria exist when the extrinsic variable is not correlated with some real shock. However, they do find evidence that sunspots matter, taking the form of common past experience that influences agents behavior even when the real shock (correlated with sunspots) has been removed. This is a combination of hysteresis, i.e. the lagging of an effect behind its cause, and sunspots. What the experiment in section 3 tries to figure out is the existence of hysteresis without sunspots. In this case discriminatory tastes are the key variable that has real effects and that is withdrawn, while there is no extrinsic signal that drives the behavior of agents after the real shock disappears.

5 Results

First of all, from the quizzes it is possible to infer that subjects have an excellent comprehension of the game, given that the average ratio of correct answers is 0.95.

The seven sessions provide interesting although somehow contradicting evidence.[11] In Figures 1-7 it is possible to have quick summary of the average offer made by Reds and Blue players.

It is immediately evident that the prediction that Reds should always bid 10 regardless of their beliefs and regardless of the choices of the Blues finds unambiguous support. Only in one case (see Figure 5A) it is possible to identify a systematic departure from the predicted behavior. What happens in Session

by Shubik (1971), where a dollar is awarded to the highest bid. Since the expenditures are sunk, it would be rational to increase a bet whenever doing so increases the expected return more than the amount of the additional bet. There is no stable equilibrium (at least in pure strategies) as long as the endowment of each player exceeds the value of the prize. When one bid exceeds the value of the prize, the motivation of the remaining bidders changes from a desire to maximize returns to one of minimizing losses. Thus, the question transforms from “How much can I win?” to “How do I keep from losing?” and escalation is easily observed, like in the classroom experiments described by Murnigham (2001).

[11] The two parts of each session, before and after the switch of the colour label, are called "A" and "B", respectively. For instance, 1A stands for the first half of the first session.

5A, one of the sessions where players were matched once and for all, is that one of the two Reds alternates a bid of 10 cents with a very low bid (2-3 cents) in the first trials. Such a behavior seems to be a signal to the other red player that it would have been more convenient to coordinate on a low bid. After a few attempts without any positive feedback from the other Red, also this player starts bidding 10 cents regularly.

The behavior of the Blue is much less straightforward to predict, since it crucially depends on players' beliefs. Not surprisingly, very different patterns emerge from the experiment. For instance, Session 5A perfectly replicates the Self-Confirming Equilibrium driven by wrong beliefs described in Section 3.3. After having learned rather quickly that the fraction of crazy computers at the beginning of the experiment is quite high (it is 87%, indeed) Blues start bidding 1 cent, which is the optimal choice given a sufficiently low chance of getting the prize. From that point onwards, both Blues always offer 1 cent and never receive evidence that contradicts the goodness of their choice, even though after the 10th trial the fraction of crazy computers is sharply and steadily decreased to 13%. Their lower bids make them less likely to win, leading to unequal outcomes that are consistent with wrong expectations that they were less likely to get the prize. It is probably not by chance that such a Self Confirming Equilibrium emerges in one of the sessions where the matching among players happens once and for all at the beginning of the experiment, i.e. where the possibility of learning from other players is lower.

A completely different pattern emerges in Session 2A (see Figure 2A), where Blues learn almost perfectly the true state of nature, both during the first trials, when the fraction of crazy computers is 87%, and after the 10th trial when it is 13%. This is evidence against the aforementioned Self Confirming Equilibrium.

Having highlighted the two polar cases emerging from the experiment, it is left to the interested reader going through the patterns session by session. There are sessions in which Blues learn to some degree that the fraction of crazy computers is lower at the end. Other sessions in which this clearly does not happen. There are also cases in which Blues do not learn that at the beginning the fraction of crazy computers is high. Given this wide range of patterns, it seems more useful to stress some regularities rather than going through all the details.

First of all, how often do subject learn that the fraction of crazy computer was much lower at the end of the session? The Wilcoxon signed-ranked test provides useful insights. By comparing the offers made by Blues during the first 10 trials with those made during the last 10 trials of every session, the Wilcoxon test rejects the null hypothesis that the offers in the two different periods are drawn from the same distribution in 8 out of 14 cases (See Table 2). In one of such cases (Session 5A) we have already seen that it is because Blues are stacked in the Self Confirming Equilibrium with wrong beliefs, and therefore their offers in the end are systematically lower than their offers at the beginning. In the other seven cases, the test rejects the null hypothesis because offers were systematically higher in the end. In other words, there is some kind

Table 2:
Wilcoxon signed-ranked test: Blues' offers (first 10 trials Vs last 10 trials).
Null Hypothesis: obs drawn from the same distribution

session	FIRST HALF	SECOND HALF
1	P>|z|: 0.1333	P>|z|: 0.4913
2	P>|z|: 0.0181	P>|z|: 0.0000
3	P>|z|: 0.0110	P>|z|: 0.0039
4	P>|z|: 0.0003	P>|z|: 0.0002
5	P>|z|: 0.0028	P>|z|: 0.2355
6	P>|z|: 0.4248	P>|z|: 0.0029
7	P>|z|: 0.1774	P>|z|: 0.2727

of learning of the true state of nature, i.e. that the actual fraction of crazy computers is lower.

Second, when it takes place, learning contradicts the existence of the Self Confirming Equilibrium driven by wrong beliefs, but its importance should not be overemphasized for two reasons:

1) a Self-Confirming Equilibrium with wrong beliefs is something that can happen but cannot be expected to emerge regularly on an aggregate basis;

2) the Self-Confirming Equilibrium is a limit situation. However, even when learning happens, persistence of discrimination can still be observed.

Looking at the distribution of prizes, the true effect of discrimination when the fraction of crazy computers has been decreased to 13% should be that, on average, Reds should get about 34 of the 60 cents available. Even focussing on the last five repetitions of the seven sessions where learning emerged in a cleaner way, it emerges that on average Blues gets 22.35 in each auction, instead of 26.1, which is the expected gain of the blues given 13% of crazy computers. Even more striking, in 44% of the cases at least one Blue is not awarded the prize. Therefore, discrimination seems to have persistent effect, even when players are not stacked in the Self-Confirming Equilibrium driven by wrong beliefs.

Third, offers of Blues are systematically higher in the second part of the sessions, i.e. for the players that have the opportunity of playing first as Reds. In other words, there are framing effects, and not only at first sight, as shown by the results of the regression presented below. Although it can introduce a non-desirable confounding factor, it is worth stressing that switching the color label in the middle of each session has also some pros, and in particular that it prevents those who are assigned to the Blue population at the beginning of each session from feeling tempted to hinder the experiment.

Apart from the Self-Confirming Equilibrium that emerges in one of the sessions where matching of subjects happens once and for all, there are no specific features of the results linked to the different setting that characterizes some sessions, either related to the length (25 Vs. 30 trials) or related to the possibility of learning also when making a low bid.

Given the design of the experiment, and in particular that every subject is exposed to the whole set of parameter changes, a regression analysis of the data is certainly informative. The limitations imposed by the very low number of independent observations prevent inference from being reliable. However, the interpretation of a regression as a conditional expectation function is not at all affected by the low number of independent observations and sheds more light on the data. Having in mind to explain the behavior of the Blues in a multivariate framework where players' bid is the dependent variable, all the regressors apart from the constant are interacted with a dummy variable taking a value equal to one for all the observations in which the players is Blue. In this way, the constant simply represents the average bid of the Reds. The coefficients of the other variables should instead be interpreted as deviations of the Blues from the average offer of the Reds. The regression, summarized in Table 3, confirms the findings above.

There is evidence that learning take place, as the inverse relation between bids and the actual fraction of crazy computers shows, although the effect of the decrease of the latter from 87% to 13% is not dramatic, accounting for a higher bid by 0.6 cents only. There are however other effects on top of this: the coefficients of the linear and quadratic trend indicates a U-shaped pattern of bids, which increase by roughly another cent between the minimum in the 18th trial and the end in the 30th trial. Bids also turn out to be much higher (by 3.1 cents) when in the previous trial an unbiased computer is discovered. However, it would be misleading to attribute a causal interpretation to this relation. Likely, the coefficient simply captures the correlation between a greater share of high bids when there are less crazy computers, but it can also be that higher bids allows to discover the type of computer and not viceversa.

There is also evidence that learning, although substantial, is much less than perfect. The dummy variable "Blue" captures that their offers are on average 2.6 cents lower than those of the Reds for no other reason than simply being Blue. However, a perfect learning would imply that the color lable should not matter given that only 13% of the computers are crazy from the 11th trial and since the actual fraction of crazy computers is controlled for. Therefore, a coefficients of -2.645 indicates that the differences between Reds and Blues are far from vanishing and hence learning is far from being perfect.

The last effect is much lower (by 1.761 cents) for the subjects that play as Blue in the second part (what has been called part "B") of the session. This points towards the existence of relevant framing effects.

Finally, personal characteristics are not very relevant in shaping subjects' behavior. Although the size of the coefficients is not negligible in some cases, it is worth stressing that the support of the regressors is rather small, and therefore personal characteristics do not explain much of the variation of bids. Particularly interesting is the fact that males and females behave in a similar manner, which implies that the setting of experiment managed to neutralize the role of a characteristic that, out of the laboratory, could be highly correlated with the behavior that the experiment was aimed to test, i.e. expectations of being discriminated against.

Table 3:
Multivariate analysis of the bids

VARIABLE	COEFF
Constant (Average offer of the Reds)	9.919
CHARACTERISTICS OF THE BLUES:	
Blue	-2.645
Blue in the second half of the experiment	1.761
Actual fraction of crazy computer	-0.008
Linear trend	-0.314
Quadratic trend	0.009
Observed an unbiased computer in the previous trial	3.103
Female	-0.222
Political orientation (0=left; 5 right)	0.091
Religious orientation (0=non believer; 2=churchgoer)	-0.327
Age	-0.434

6 Conclusions

This paper is aimed at testing the predictions of a model that explores the role of workers' expectations of being discriminated against as an original explanation for the puzzling long-run persistence of observed discrimination against some minorities in the labor market. The model, presented in Section 2, provides a theoretical framework based on a two-stage game of incomplete information where preferences and beliefs of both sides of the labor market matter. In every stage game two workers, one of whom is a minority worker, are drawn from their *ex ante* identical populations and randomly matched with one employer.[12] At the end of the first period the employer promotes one (and only one) worker after having observed the output they have produced, which is one to one related with effort. Promotions also depend on employer's type, unknown to the workers, which captures the possible disutility of promoting a minority worker. The importance of workers' expectations can be appreciated by comparing the distribution of promotions across populations that arises when minority workers overestimate the percentage of employers characterized by tastes for discrimination with a situation in which such beliefs are correct *ceteris paribus.* This difference becomes crystal clear when there are actually only employers who do not discriminate against the minority either directly or statistically. Even in this circumstance unequal outcomes may emerge, caused by wrong beliefs of being discriminated against that are self-confirming. Minority groups who expect of being discriminated against exert a lower effort on average, because of a lower expected return. This induces a lower observed percentage of promotions within minority workers, which in turn is consistent with their beliefs that there are employers characterized by discriminatory tastes.

[12] What distinguishes the population of minority workers is an observable characteristic not related to their productivity (e.g. race, gender).

The experiment replicates the model using a game where participants are randomly divided into two populations: red and blue. In every trial each participant has an endowment of 10 Euro cents and can decide how many cents to allocate as a sort of lottery ticket to get a prize worth 60 Euro cents. Bets are not given back to the players, neither to the winner nor to the loser, making the game equivalent to an all-pay auction. In every lottery there are four participants, two from each population. The prize is awarded to the higher bid and it is split if bids are equal, unless the opponents are assigned to a crazy computer which instead awards the prize to the red player regardless of the bids. Subjects do not know the probability of facing a crazy computer.

This experiment follows an old version that failed to provide convincing evidence because it didn't manage to separate beliefs of subjects belonging to the two different populations. The main difference in this new version is that there are four players instead of two competing for the prize in each auction. This change of the setting is crucial because it make the behavior of the Reds unaffected by the type of computer, since it induces them to bid always 10 cents. Therefore, the fraction of crazy computers, and consequently beliefs of Reds about it when such a fraction is unknown, do not play any role in shaping their behavior. This trick directly implements a behavior of the Reds equivalent to what would be observed in the old version when the red player has correct beliefs about a sufficiently low fraction of crazy computers. In this way, what is left to analyze is just the behavior of the Blues, which instead crucially depends on their beliefs when the fraction of crazy computers is unknown.

The mechanism underlying the Self-Confirming Equilibrium driven by wrong beliefs in the theoretical model is tested comparing the behavior of advantaged (red) and disadvantaged (blue) subjects, once the fraction of crazy computer, kept equal to 87% for the first ten trials, is suddenly decreased to 13% from that moment onwards.

The results of the experiments show that not surprisingly the behavior of Reds conform extremely closely to the theoretical predictions. Very different patterns emerge instead as far as the Blues, not surprisingly even in this case since their behavior crucially depends on their beliefs about the unknown fraction of crazy computers. One session perfectly replicates the Self-Confirming Equilibrium driven by wrong beliefs: after having learned rather quickly that the fraction of crazy computers at the beginning of the experiment is quite high (it is 87%, indeed) Blues start bidding 1 cent, which is the optimal choice given a sufficiently low chance of getting the prize. From that point onwards, both Blues always offer 1 cent and never receive evidence that contradicts the goodness of their choice, even though after the 10th trial the fraction of crazy computers is sharply and steadily decreased to 13%. Their lower bids make them less likely to win, leading to unequal outcomes that are consistent with wrong expectations that they were less likely to get the prize. In half of the sessions the Wilcoxon signed-rank test rejects that learning takes place, while the others are characterized by a significant although not perfect learning of the underlying percentage of crazy computers. What is important to stress is that, even where learning takes place, the Blues do not reach a situation of balanced outcomes. For instance, focussing on the last five repetitions of the seven sessions where learning emerged in a cleaner way, it emerges in 44% of the cases at least one Blue is not awarded the prize. Therefore, discrimination seems to have persistent effect, even when players are not stacked in the Self-Confirming Equilibrium driven by wrong beliefs, and it seems possible to conclude that expectations are likely to increase the probability that observed discrimination lags behind its causes. One session perfectly replicates the Self-Confirming Equilibrium driven by wrong beliefs: after having learned rather quickly that the fraction of crazy computers at the beginning of the experiment is quite high (it is 87%, indeed) Blues start bidding 1 cent, which is the optimal choice given a sufficiently low chance of getting the prize. From that point onwards, both Blues always offer 1 cent and never receive evidence that contradicts the goodness of their choice, even though after the 10th trial the fraction of crazy computers is sharply and steadily decreased to 13%. Their lower bids make them less likely to win, leading to unequal outcomes that are consistent with wrong expectations that they were less likely to get the prize. In half of the sessions the Wilcoxon signed-rank test rejects that learning takes place, while the others are characterized by a significant although not perfect learning of the underlying percentage of crazy computers. What is important to stress is that, even where learning takes place, the Blues do not reach a situation of balanced outcomes. For instance, focussing on the last five repetitions of the seven sessions where learning emerged in a cleaner way, it emerges in 44% of the cases at least one Blue is not awarded the prize. Therefore, discrimination seems to have persistent effect, even when players are not stacked in the Self-Confirming Equilibrium driven by wrong beliefs, and it seems possible to conclude that expectations are likely to increase the probability that observed discrimination lags behind its causes.

Appendix: Instructions

The Experiment - part A (shown at the beginning)

The experiment will last approximately 60 minutes, but the actual length depends on the speed of the slowest participant. The experiment is composed by two quizzes, two stages and a questionnaire.

Numbers during the experiment represent Euro cents. Your final earnings will be the sum of all the Euro cents you earned throughout the experiment. Earnings depend on your choices as well as on the choices of your opponents during the game that will start in a few minutes.

The game consists of an auction, in which you have to bid in order to get a prize. The game will be repeated 50 times.

At the beginning of the experiment an algorithm will assign to every player a color label (red or blue) that will be effective for half of the experiment, i.e. 25 repetitions. The two populations (Red and Blue) will be of equal size. In every repetition of the game your opponents will be three anonymous players: one drawn from the same population as yours, the other two from the other population (i.e. if you belongs to the red population you will always play against another Red and two Blues and vice versa). Before every repetition, an algorithm will randomly create groups of four players: therefore, your opponents will change from repetition to repetition.

Let us forget about the color for the moment, we will talk about it again in details in a few minutes. What follows are the rules of the game as if the color was not relevant:

In every repetition of the game you will be endowed with 10 Euro cents. You have to decide how much to bid (from 1 to 10 cents) in order to win a prize worth 60 Euro cents.

- The higher bid wins the prize.
- If bids are equal, the prize is equally split.
- Bids are not given back, neither to the winner(s) nor to the loser(s).

Your earnings in every repetition of the game depends on two factors:
1) The prize: 60 cents if you only are awarded the prize, 30 cents if the prize is split with another player, 20 cents if split with other two players, 15 cents if all players made the same offer, and finally 0 cents if some your opponents is awarded the prize;
2) How much of your endowment you did not bid.

N.B. You cannot save and transfer money from one repetition of the game to another. If you bid less than 10 cents the amount left will enter your earnings but in the following repetition you will start again with 10 cents.

The Experiment - part B (shown after Quiz 1)

Another kind of computers is part of the game. These computers, which are called "crazy computers," always award the prize among the red players regardless of whether some Blues made higher bids. If the outcome of the auction is decided by a crazy computer:

- The higher bid **among the Reds** wins the prize.
- If bids of the Reds are equal, the prize is equally split.
- Bids are not given back, neither to the winner(s) nor to the losers.

Note that the computers we are talking about ("unbiased" vs. "crazy") are not the computers you have in front of you. The server computer has been programmed to receive the data from the client computers (i.e. to receive the bids that you enter in the PCs in front of you). In every repetition of the game, each group of four players is randomly associated with a partition of the server that can correspond to an unbiased computer (which assigns the prize to the higher bid(s)) or to a crazy computer (which never awards the prize to the Blues).

NB: In every repetition of the game, the bids of each group of players are randomly assigned to a partition of the server computer. Hence, being assigned to a crazy or to an unbiased computer during one repetition of the game does not provide additional information about the type of computer you will face in the following repetition. From this point of view it is like starting from the beginning at every repetition. You only have to bear in mind that each time there is a percentage of partitions of the server that represent crazy computers.

How many are the crazy computers? The percentage of crazy computers can vary from 0% to 100% during the experiment and you will know such fraction.

The introduction of crazy computers creates different conditions for Reds and Blues. To ensure that all participants have the same earning opportunity, the color label will be switched after 25 repetitions. Hence, if you are now assigned to the Blue population, you will be Red in the second half of the experiment. At the beginning of every repetition the display will remind you the color of your population.

References

[1] Anderson, D.M. and M.J. Haupert (1999): "Employment and Statistical Discrimination: a Hands-On Experiment," *Journal of Economics*, 25:1, 85-102.

[2] Anderson, L.R., R.G. Fryer and C.A. Holt (2002): "Experimental Studies of Discrimination," in *Handbook on Economics of Discrimination*, ed. by W. Rogers, forthcoming

[3] Anderson, S.P., J.K. Goeree and C.A. Holt (1998): "Rent Seeking with Bounded Rationality: An Analysis of the All Pay Auction," *Journal of Political Economy*, 106:4, 828-853

[4] Arrow, K.J. (1973): "The Theory of Discrimination," in *Discrimination in Labor Markets* ed. by O. Ashenfelter and A. Rees. Princeton: Princeton University Press, 3-33.

[5] Baye, M.R., D. Kovenock and C.G. De Vriers (1996): "The All-Pay Auction with Complete Information," *Economic Theory*, 8, 291-305.

[6] Bull, C., A. Schotter and K. Weigelt (1987): "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, 95:1, 1-33.

[7] Davis, D.D. (1987). "Maximal Quality Selection and Discrimination in Employment," Journal of Economic Behavior and Organization, 8, 97-112.

[8] Davis, D.D. and C.A. Holt (1993): *Experimental Economics.* Princeton: Princeton University Press.

[9] Davis, D.D. and R.J. Reilly (1994): "Do too Many Cooks Always Spoil the Stew? An Experimental Analysis of Rent-Seeking and the Role of a Strategic Buyer," Public Choice, 95, 89-115.

[10] Duffy, J. and J. Ochs (2003): "Cooperative Behavior and the Frequency of Social Interaction," mimeo, University of Pittsburgh.

[11] Filippin, A. (2003a): "Discrimination and Workers' Expectations," IZA Discussion Paper n.823.

[12] Filippin, A. (2003b): "Discrimination and Workers' Expectations: Experimental Evidence," IZA Discussion Paper n.824.

[13] Filippin, A. and A. Ichino (2004): "Gender Wage Gap in Expectations and Realizations," *Labour Economics,* forthcoming.

[14] Fershtman, C. and U. Gneezy (2001): "Some Experimental Evidence on the Evolution of Discrimination, Cooperation and Perceptions of Fairness," Economic Journal, 481, 679-703.

[15] Fischbacher, U. (2002): "zTree Experimenter's Manual," mimeo, University of Zurich.

[16] Fryer, R.G., J.K. Goeree and C.A. Holt (2002): "An Experimental Test of Statistical Discrimination" mimeo, The University of Virginia.

[17] Hargreaves-Heap, S. and Y. Varoufakis (2002): "Some Experimental Evidence on the Evolution of Discrimination, Cooperation and Perceptions of Fairness," Economic Journal, 481, 679-703.

[18] Lazear, E.P and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," Journal of Political Economy, 89:5, 841-64.

[19] Lohm, S. (2000) "Stereotypes, Sex Discrimination and Paranoia - an experimental study," Scandinavian WP in Economics 2000:1.

[20] Marimon, R., S. Spear and S. Sunder (1993). "Expectationally-Driven Market Volatility: Experimental Evidence," Journal of Economic Theory, 61, 74-103.

[21] Murningham, J.K. (2001): "A Very Extreme Case of the Dollar Auction," mimeo, Kellogg Graduate School of Management, Northwestern University.

[22] O'Keeffe, M., W. K. Viscusi, and R. J. Zeckhauser (1984): "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 2:1, 27-56.

[23] PHELPS, E.S. (1972): "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62:4, 659-661.

[24] ROTH, A.E. (1994). "Let's Keep the Con out of Experimental Econ.: A Methodological Note," *Empirical Economics*, 19, 99-109.

[25] SCHOTTER, A. AND K. WEIGELT (1992): "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results," *Quarterly Journal of Economics*, 107, 511-39.

[26] SHUBIK, M. (1971): "The Dollar Auction Game: A Paradox in Non-Cooperative Behavior and Escalation," Journal of Conflict Resolution, 15109-111.

[27] WOODFORD, M. (1990). "Learning to Believe in Sunspots," Econometrica, 58:2, 277-308.

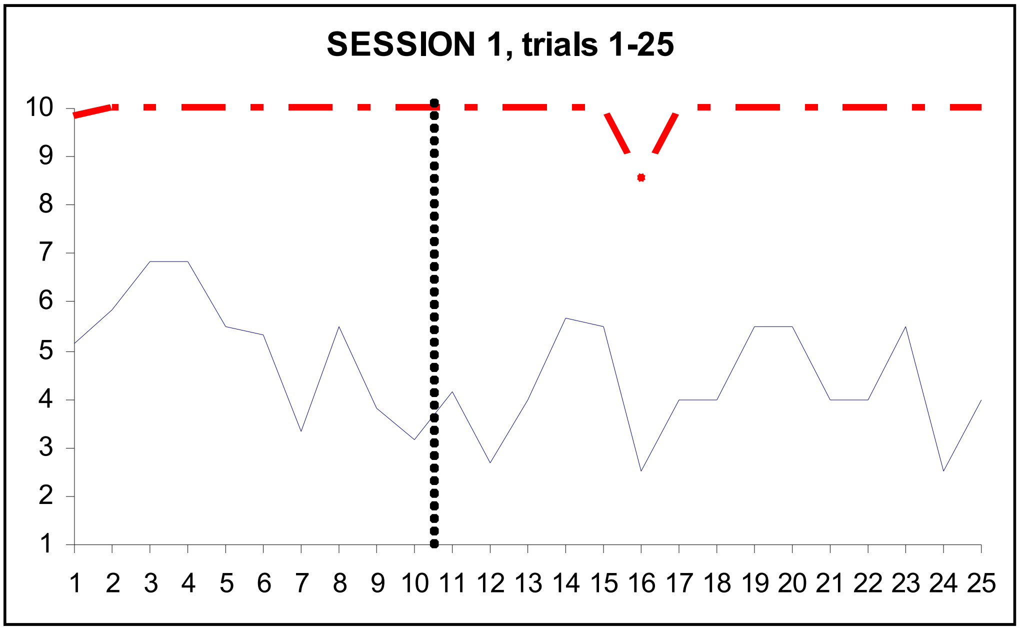

Figure 1a: Session 1, trials 1-25. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.1333

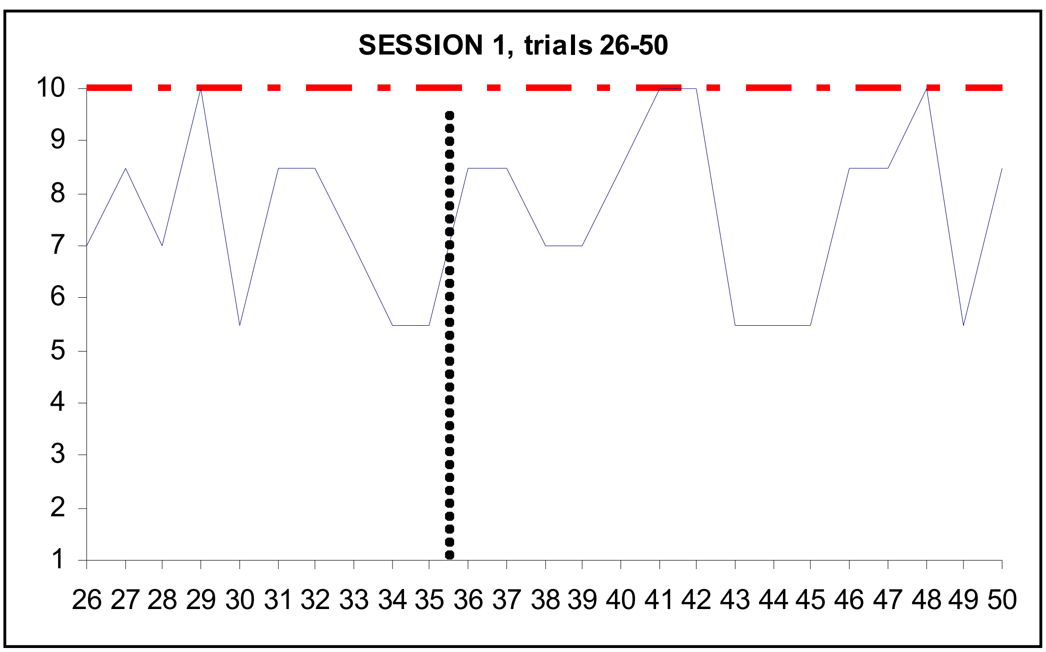

Figure 1b: Session 1, trials 26-50. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.4913

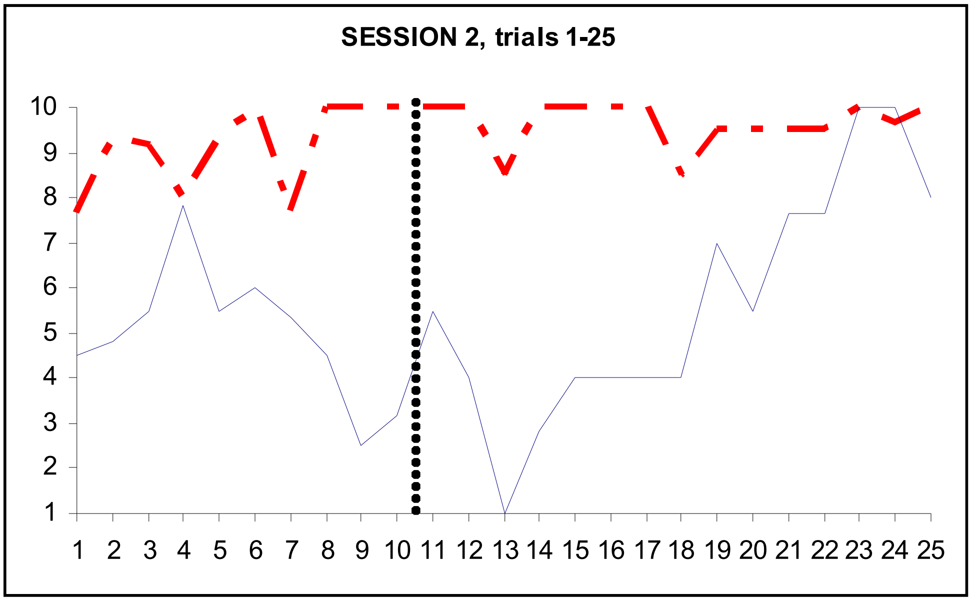

Figure 2a: Session 2, trials 1-25. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0181

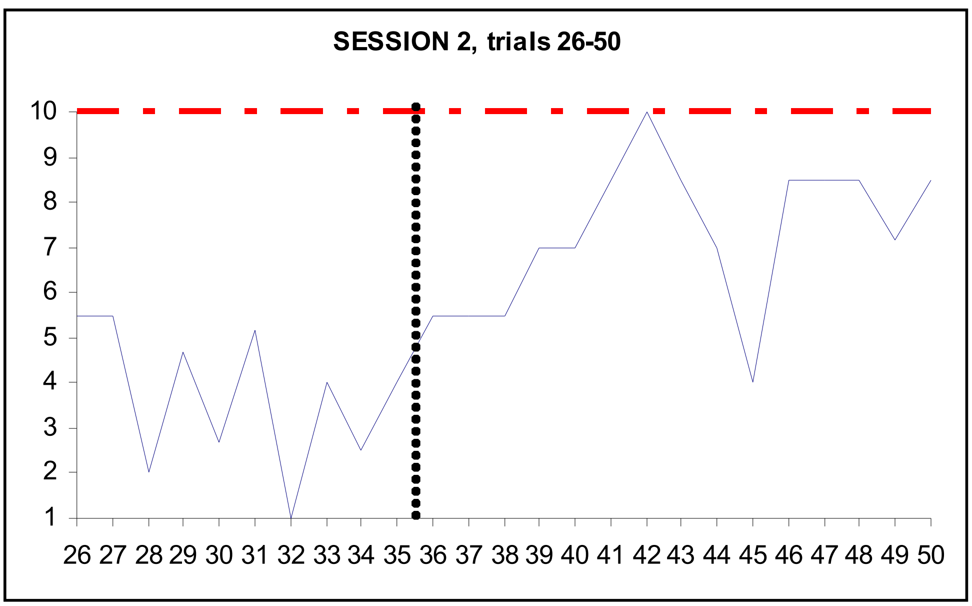

Figure 2b: Session 2, trials 26-50. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0000

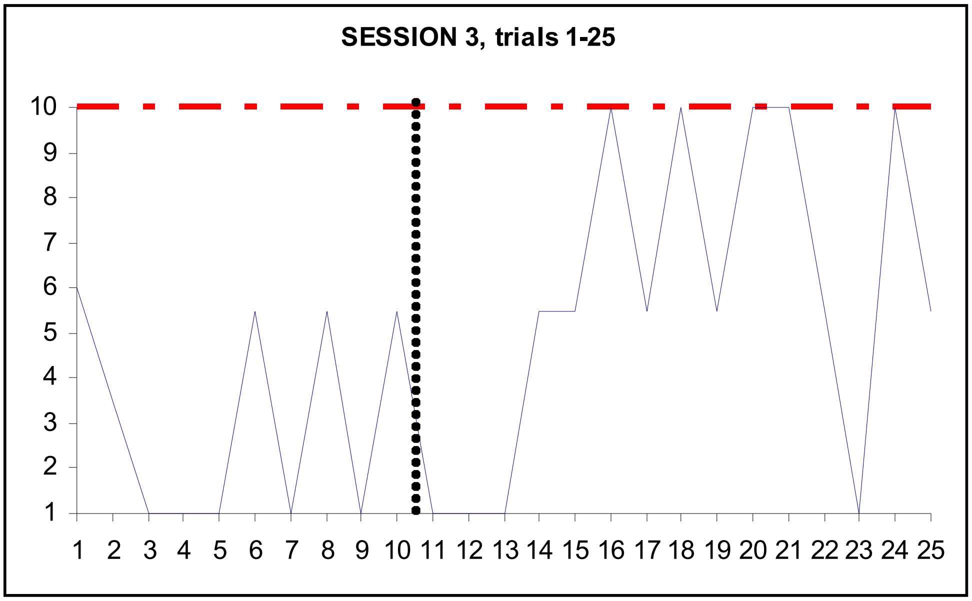

Figure 3a: Session 3, trials 1-25. Number of participants: 4
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0110

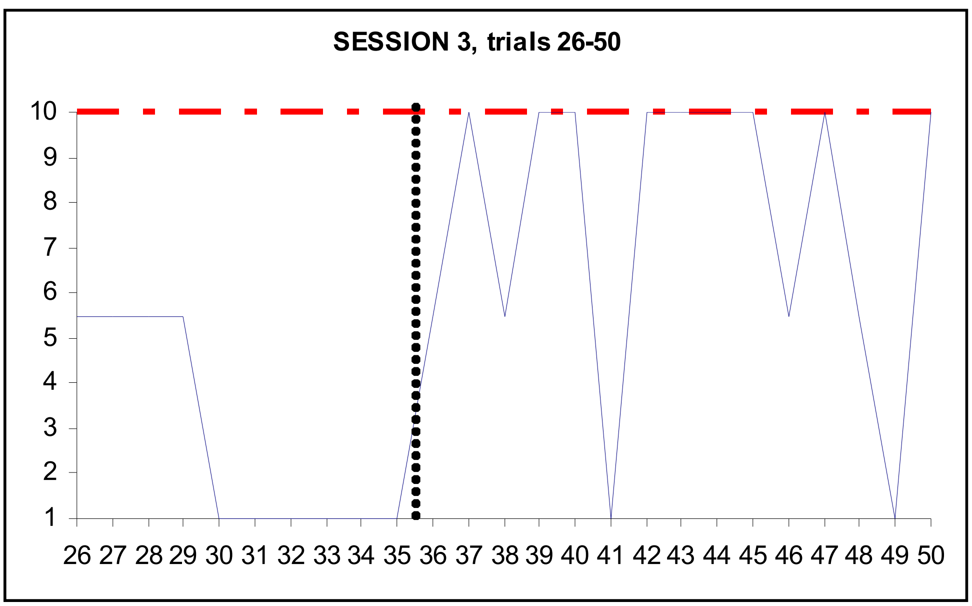

Figure 3b: Session 3, trials 26-50. Number of participants: 4
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0039

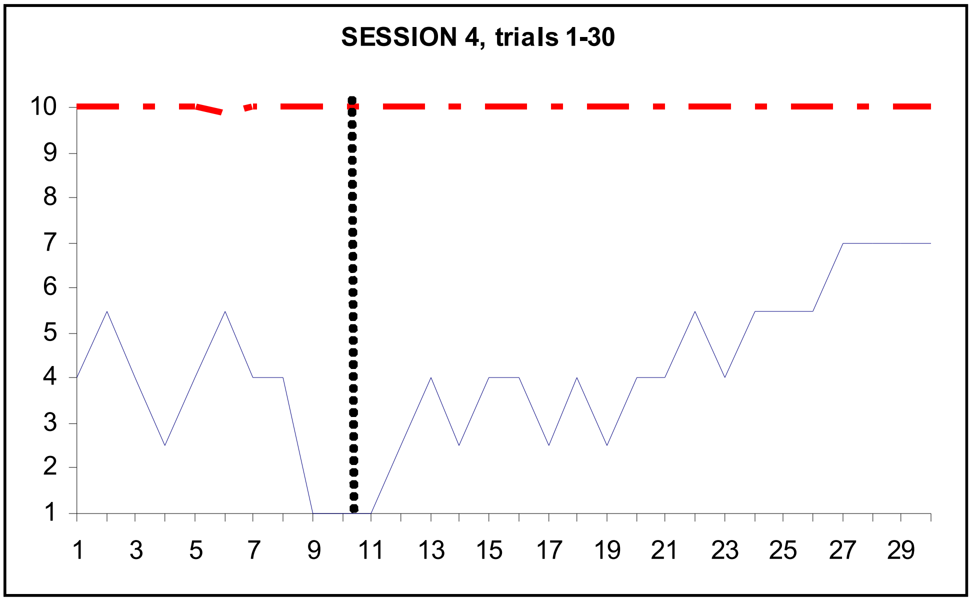

Figure 4a: Session 4, trials 1-30. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0003

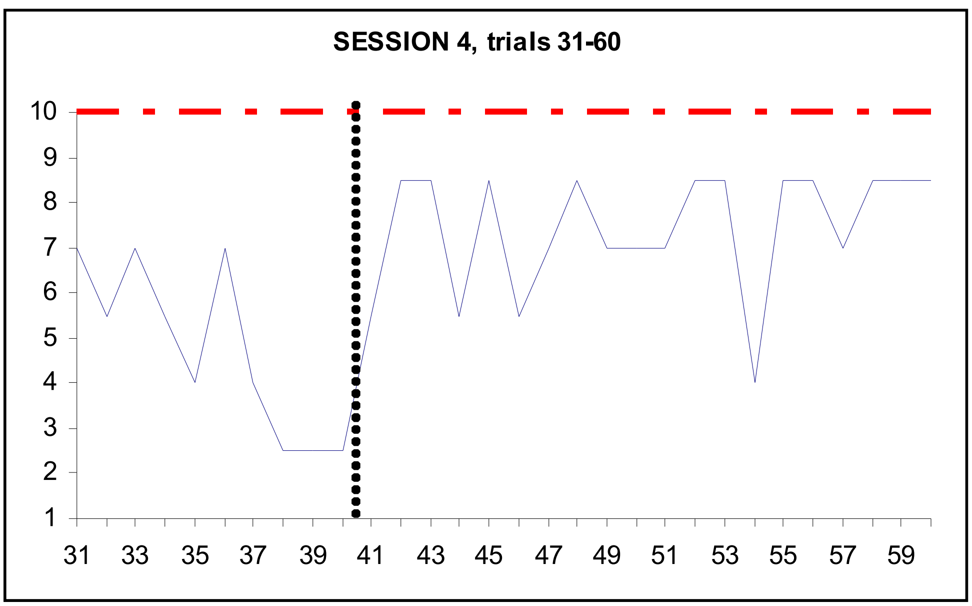

Figure 4b: Session 4, trials 31-60. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0002

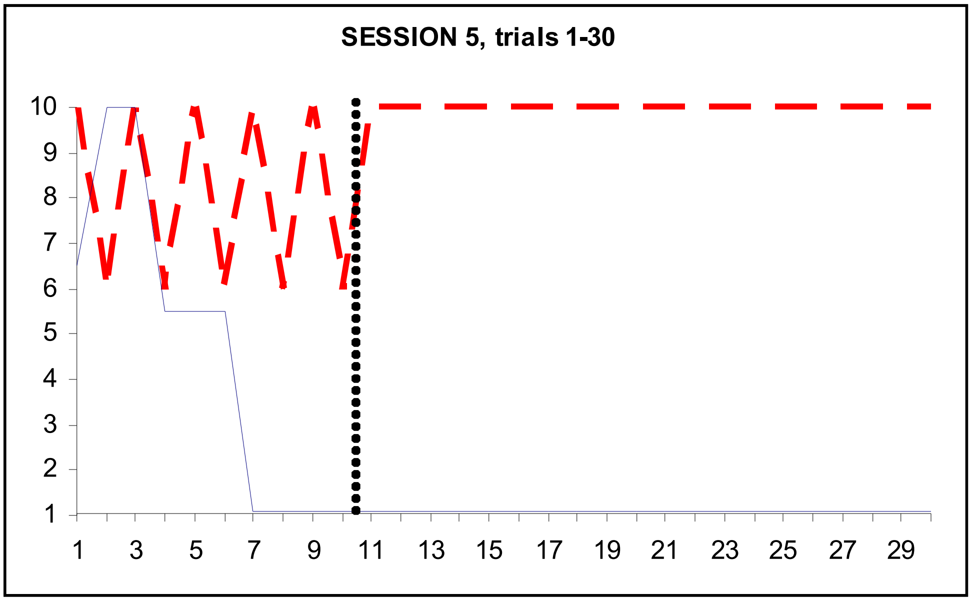

Figure 5a: Session 5, trials 1-30. Number of participants: 4
Average offer of REDS (dashed line) and BLUES (solid line

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0028

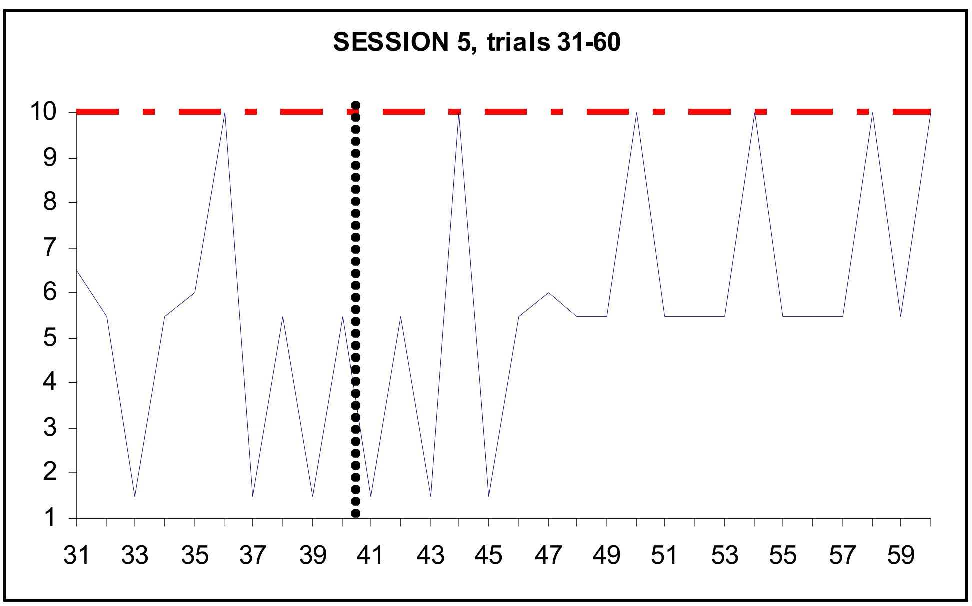

Figure 5b: Session 5, trials 31-60. Number of participants: 4
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.2355

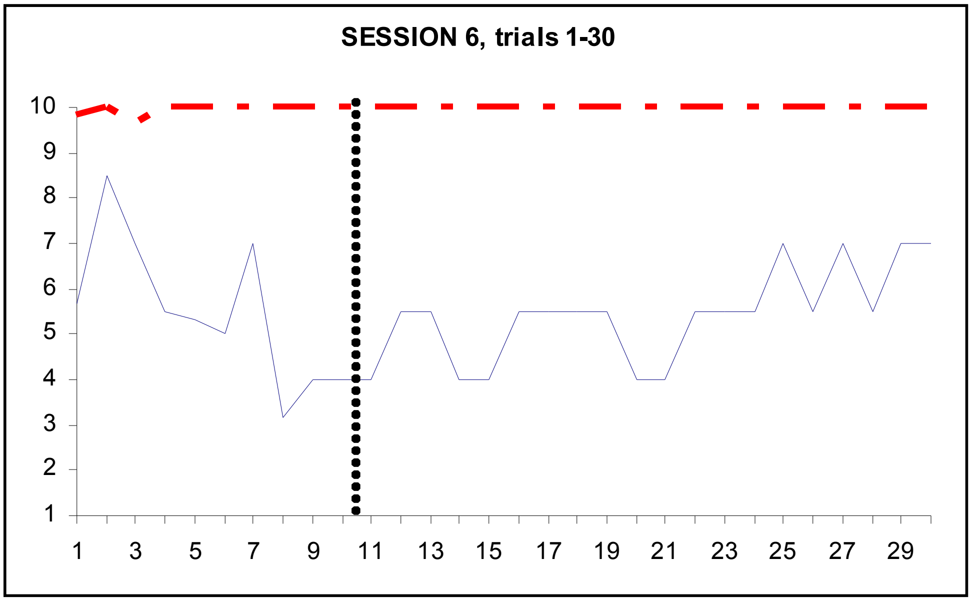

Figure 6a: Session 6, trials 1-30. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.4248

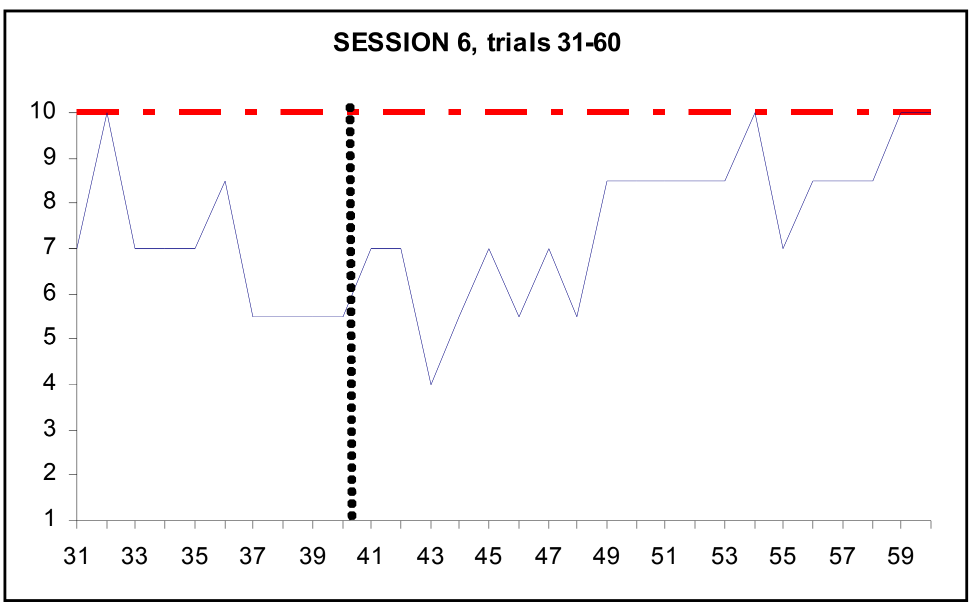

Figure 6b: Session 6, trials 31-60. Number of participants: 12
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.0029

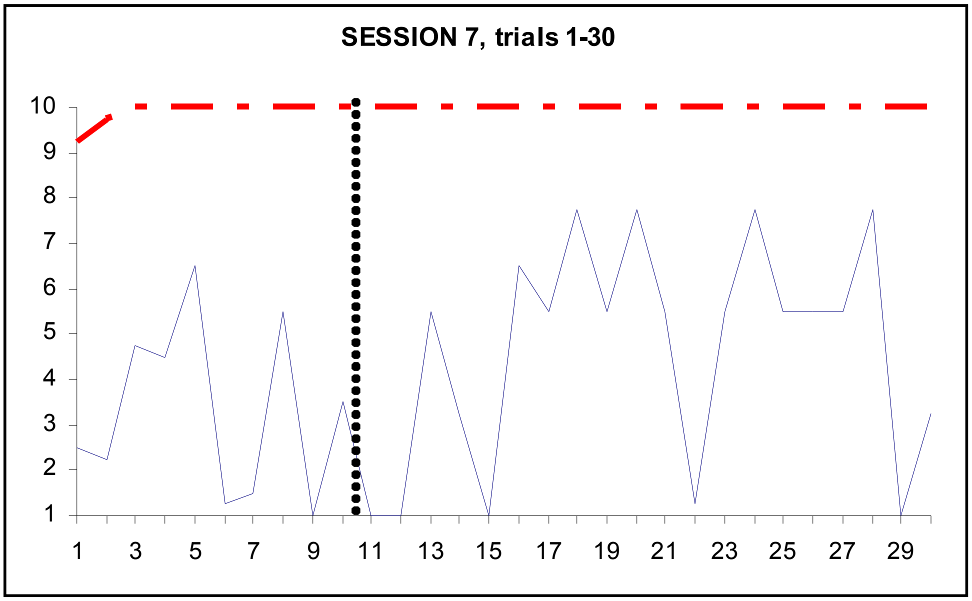

Figure 7a: Session 7, trials 1-30. Number of participants: 8
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.1774

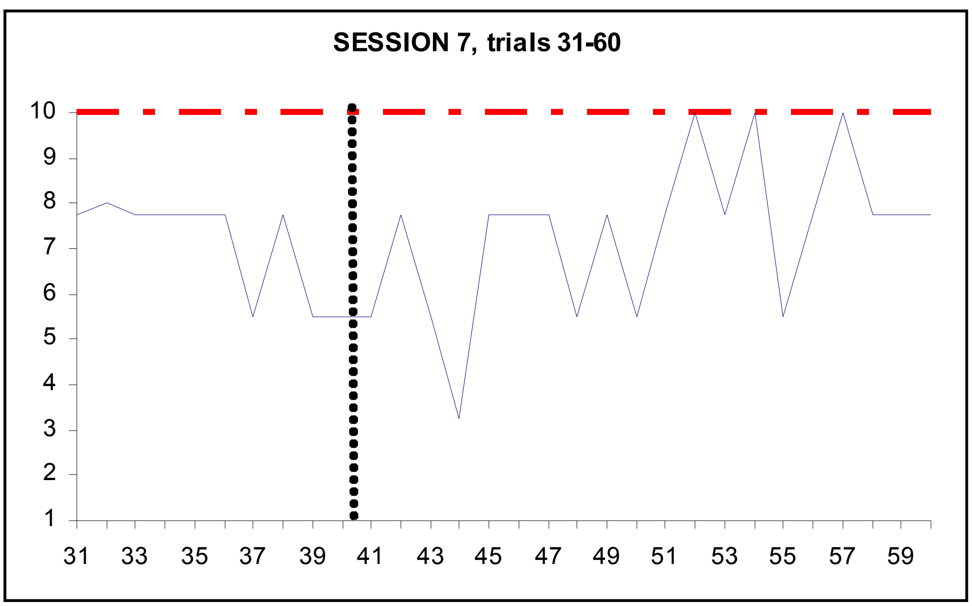

Figure 7b: Session 7, trials 31-60. Number of participants: 8
Average offer of REDS (dashed line) and BLUES (solid line)

Wilcoxon signed-rank test (Blues: first 10 trials Vs last 10 trials): P>|z|=0.2727