JOURNAL OF PATTERN RECOGNITION RESEARCH 1 (2009) 133-151

Received Feb 06, 2009. Accepted May 31, 2009.

Behavioral Pattern Learning Models for Decision Making in Games

Yaron Lahav YARON@LAHAVCONSULTING.COM
Lahav Consulting, LLC
6701 Democracy Blvd. Suite 300, Bethesda, MD, 20871, USA

Abstract

This paper presents and tests a new learning model of boundedly rational players interacting with nature. According to the model, players look back on the history of choices made by nature and search for patterns. After updating their beliefs about the likelihood of the reappearance of each pattern, the players choose the best response to more likely patterns. A probabilistic decision rule governs the tendency of the players to choose (or not to choose) certain actions, along with the influence of the payoff for each action. The parameters of the model are estimated using data from a new experiment, and the model is compared with other existing learning models.

1. Introduction

The study of individual decision making, strategic behavior, and learning in games has advanced significantly in the last two decades, in part as a result of the increasing use of the experimental approach. This approach enabled the use of decision making data that is not confounded with other factors. As soon as the experimental data revealed that at times, players do not play in accordance with the traditional theories (such as expected utility maximization or Nash equilibrium), learning models with the assumption of bounded rationality were proposed. These models apply earlier psychological theories [9], use some economic principles [5, 13, 21] or combine the two approaches [3, 23, 26].

In this paper we present a new behavioral learning model. This model describes the ability of individuals to recognize patterns in someone else's behavior and use this ability as a "learning technique" for predicting the other person's future outcomes. This technique is evaluated and tested in an environment where individuals need to guess future outcomes that are exogenous to the individuals own decision. We allow players to search for deterministic patterns when their opponent does not play strategically . Assume two players play "matching pennies" several times. Also assume that one of the players makes her choices by using some kind of a pattern by choosing "heads" in even periods and "tails" in odd ones (choosing "heads" followed by "tails", followed by "heads" again and so on). It is reasonable to assume that the other player will recognize the alternating pattern at some point in time, and take advantage. Using the learning rules above, her opponent might never recognize the pattern. [1]

The study of the way people learn to predict future outcomes is of interest at both the individual and the aggregate levels. It is relevant in particular when the outcome is a realization of an aggregate variable, although the individual is assumed not to have the ability to influence the outcome. A good example is the performance of asset markets. Although the single investor has no influence on the future performance of these markets, future prices are influenced by the sum of all individuals' investment decisions. By characterizing individuals' behavior according to their learning methods and ability, an observer may be able to use these predictions to forecast outcomes in those markets. It is argued [2, 15, 16] that traders' sentiments are important in predicting future price changes in

1 This does not mean that the model presented is not appropriate to predict the learning process of individuals that play against strategic players. We did not test our model in that setting, and therefore cannot say much about it empirically.

© 2009 JPRR. All rights reserved. Permissions to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or to republish, requires a fee and/or special permission from JPRR.

WWW.JPRR.ORG

asset markets. So knowing how traders form their beliefs regarding the future performance of asset markets can help us to predict the future performance of these markets.

We present two versions of a probabilistic model (one version is a two-parameter model and the other is a four-parameter model) that uses this approach to estimate human players' decisions. Players use historical observations to recognize patterns. They do so because they believe that this is the only way to guess future behavior. The human players, however, do not have unlimited processing and memory abilities. Therefore, the search for the pattern is limited to those processing and memory abilities that lie within individuals' capabilities. A difference in cognition is allowed when recognizing patterns that occurred in the distant versus the recent past, and there may be different reactions for long patterns versus short ones.

The model allows the payoff matrix to influence an individual's choices. Intuitively, if a change in the payoff matrix changes the player's decision, it would contradict the argument that the player looks for patterns in nature's behavior. The process of pattern recognition, at least in principle, should not be affected by the payoff associated with different actions. In fact, it is not necessarily a contradiction to the existence of a pattern recognition process. It means that the process is subject to distortion. The payoff matrix determines the players' tendency towards a certain action, but not their beliefs about nature's choices (In other words, the less certain the player is about the probability distribution of nature's future choices, the more the player is influenced by the reward of choosing each action and achieving a match, and vice versa).

Indeed, a player may believe her opponent uses a pattern in behavior, even if her opponent actually does not do so. Margolis [19] also argues that individuals recognize patterns even when patterns do not exist. Sonsino [26] presents a study in which players play strategically most of the time, but may recognize patterns in their opponents' strategies from time to time. When this happens, they abandon their strategic behavior and best respond to the next move in the pattern as they remember it, the one they believe will occur according to the pattern. It should be mentioned, that Sonsino does not present a model of pattern recognition in which individuals recognize patterns as a way of learning. In his model, the recognition of a certain pattern occurs only every now and then, and the best response to a realized pattern ends up with the immediate break of the pattern and a return to strategic behavior.

Although the use of pattern recognition as a learning method is not common in economic theory, it is widely used in other disciplines such as computer science, biology, statistics and psychology [14]. In biology, researchers use pattern recognition methods (sequence analysis) to locate types of genes from a long protein sequence. In computer science, models of pattern recognition are used for speech recognition (telephone directory enquiry without human assistance) or image processing (learning to recognize a fingerprint or a face image for security purposes. See [6] as an example and [8] for review). Statistical models such as Neural Networks use pattern recognition for several purposes, including data mining (see [4] for a review on neural networks).

But this implementation of pattern recognition has its origin in human behavior. In their survey on statistical pattern recognition, Jain et al [14] argue that "the best pattern recognizers in most instances are humans." They also mention previous findings that pattern recognition is critical in most human decision making tasks. Margolis [19] argues "that pattern recognition is central to thinking is a familiar idea. Everyone who writes on these matters discusses - in some way or another - patterns for the most elaborate computerized character recognizer cannot yet match a bright six-year-old in recognizing letters of the alphabet from many different fonts."

But rare uses of pattern recognition do exist in economic theory. An example is Rötheli [22], who argues that people use pattern recognition when trying to guess future events. He uses pattern recognition as an updating rule in a learning model that assigns probabilities to each future event. Rötheli's work is similar to the model presented in this paper in that he uses a decision problem of an individual that is facing nature rather than human opponents. There are, however, two crucial differences between Rötheli's model and ours: the first is the fact that the individual uses all information regarding nature's previous moves to update her beliefs regarding nature's future moves. In our model, the individual can decide whether to use certain pieces of information or not. The second difference is in the structure of Rötheli's model: it cannot be used if nature generates more than two outcomes.

When playing against nature, the assumptions made by the human player should be different than in a game or strategic interaction. First, since nature does not have a payoff function and cannot enjoy the consequences of the actions it takes, the human player has no reason to believe that nature will make an attempt to break a pattern once a pattern is detected. When interacting with nature, the human player is more likely to think that her belief was incorrect all along if a hypothesized pattern is broken rather than that the pattern she faces was just changed by a strategic individual. Second, if nature is not human, the only reasonable way to model its outcomes is by the history of its performance. Third, by assuming that nature's moves are independent of the human player's actions, the alternative to thinking that a pattern exists is to assume that the behavior of nature is such that no order can be found.

In Section 2 we present the two-parameter version of the learning model. We begin with the learning rule, and continue with the decision rule. In Section 3 we present the four-parameter version. In Section 4 we describe the experiment and the data gathering. The results are presented in Section 5 and a summary in Section 6.

2. The Two-Parameter Model

2.1 Overview

When playing repeatedly against nature, the player first observes nature's recent moves and processes them as patterns. The player encodes the last move as a pattern consistent with nature's behavior, the last two moves as a different pattern, the last three moves as a different pattern and so on. Next, the player reviews all previous moves made by nature to find other occurrences of these recent patterns. Then, a process of belief formation occurs. Three criteria affect the process:

(a) the more a pattern occurs in the past, the more weight will be assigned to it,

(b) the timing of the occurrence of the pattern in the past has an effect on the belief formation (primacy and recency effects),

(c) the longer the pattern, the more weight will be assigned to it, provided that it appeared at least once in the past.

Beliefs about the next move of nature are formed according to what was nature's choice right after the realized pattern occurred in the past. If, for example, the pattern corresponds to the three most recent outcomes and was observed in the past, then if the player's objective is to match nature's move, the action that will be chosen by the player for the next period will be the action taken by nature in the period that followed the pattern detected in the past. Since a pattern can appear in the past more than once, and in each time the following action made by nature could be different, the player tends to choose the action with the greatest likelihood according to her beliefs.

It should be mentioned, that there is no belief accumulation in this model. In each period, the player uses only the relevant patterns, and abandons all the information gathered in the last period. Rötheli's model specifies a counter that is updated in every period. This is not the case here. If a player correctly guesses nature's choice in a certain period, then the same pattern will be used again for the next period. The player will now observe a longer pattern, and will therefore assign more weight to it. If the player's choice turned out to be the wrong one, then there is no reason to use this false pattern again.

2.2 The learning rule

Let t be a period of play, s_t be an action of the stage game at any given time t, where $s_t \in S$ (in our case: $S = \{1, 2, 3\}$) and θ be the number of elements in S. $j \in \{1, ..., J\}$ index is a pattern and n_j, denotes the length of pattern j. i denotes individual, x_t is nature's choice at time t and τ is another index of time ($1 \leq \tau \leq t$). For example, if the action space is $s = 1, 2, 3$ and nature's choices in time $t = 5, 6, 7$, were $x_5 = 1, x_6 = 3$ and $x_7 = 2$, then $j = 1, 2, 3$ and $n_j = 3$.

Next, we define an indicator variable for every pattern j with a length n_j. This variable recognizes the pattern as it ends at time t, and the same pattern in the past, ending at time τ.

$$D_{t,\tau}^{j,n_j} = \left\{ \begin{array}{ll} 1 & \textit{if } x_{t-n_j+1}, ..., x_t = x_{\tau-n_j+1}, ..., x_\tau = j \\ 0 & \textit{otherwise} \end{array} \right\} \tag{1}$$

Then, we define $y_{\tau+1}^{j,n_j}$ as the action generated by nature after pattern j with length n_j occurred in the past and ended at time τ. These variables, along with the parameters of the model (explained below) form the weight that player i assigns to each $s \in S$ at time $t+1$,$W_{i,t+1}^s$:

$$W_{i,t+1}^s = \sum_{n_j=1}^{t-1} \sum_{\tau=1}^{t-1} D_{t,\tau}^{j,n_j} C_{\tau+1}^s (\gamma_i)^{t-\tau} (\delta_i)^{n_j} \tag{2}$$

Where $C_{\tau+1}^s$ is a dummy variable such that:

$$C_{\tau+1}^s = \left\{ \begin{array}{ll} 1 & \textit{if } y_{\tau+1}^{j,n_j} = s \\ 0 & \textit{otherwise} \end{array} \right\} \tag{3}$$

This weight is cumulative over experience. The value of $W_{i,t+1}^s$ is intrinsically meaningless but is used for the formation of the beliefs. Unlike other models, this value is not accumulated over time. Instead, the process of belief formation restarts every period (notice that $W_{i,t+1}^s$ is not a function of $W_{i,t}^s$, although $W_{i,t+1}^s$ is formed on the basis of the history of nature's actions). The only change from one period to the other is the additional information faced by the player. This new information is a result of a longer history after each consecutive round.

The parameters γ and δ measure the importance of the location of the pattern in the past, and the importance of its length, respectively. For this reason they are called the recency and the length parameters. If the influence of an appearance of a pattern that occurred in the past is stronger than the influence of an appearance of the same pattern that occurred recently, then $\gamma > 1$. If the opposite is true, then $0 < \gamma < 1$. The length parameter has a different meaning, since patterns are weighted cumulatively[2]. So the higher δ is, the bigger the additional influence of longer patterns is.

It is worth mentioning, that as presented, this learning rule can be consistent with several occurrences of the same pattern in the past, followed by different moves. By using γ and δ the player uses time and length of occurrence as hierarchies to sort the patterns in a certain order when forming

2 Take, for instance, the pattern $j = 1, 2, 3$. This pattern will be weighted 3 times: once as the pattern $j = 3$, once as the pattern $j = 2, 3$ and once as the pattern $j = 1, 2, 3$.

beliefs. In this way the player can still form beliefs when a pattern occurs twice in the past, followed by different moves of nature each time.

As an example, consider a case where nature can generate two outcomes: *1* and *2*. Also assume that a single player has ended nine rounds playing against nature. The history profile of nature's outcomes is: *1,2,2,1,2,1,1,2,1*. In order to decide what nature will generate in round 10, the player first observes the smallest pattern that ends in round 9. This pattern would be $j = 1$, with a length $n_1 = 1$. Notice that $D^{1,1}_{9,7} = 1$, because the same pattern in time $\tau = t - 2$ is also 1. Since the action chosen by nature at $t = \tau + 1$ is '2', then $C^2_8 = 1$. Having all dummies with the value '1', the weight of $s = 2$, or $W^2_{i,10}$, starts to be formed.

But the pattern $j = 1$ also appeared in $\tau = t - 3$, and was followed by $s = 1$. Having $D^{2,1}_{9,6} = 1$ and $C^2_7 = 1$, the weight of $s = 1$, or $W^1_{i,10}$, also starts to be formed. The same happens for $\tau = 4$ and $\tau = 1$.

The parameter γ determines what location in the past has the most influence. If $0 < \gamma < 1$, then a recognized repetition close to $\tau = t$ will contribute the most and a recognized repetition close to $\tau = 1$ will contribute the least. The opposite is true if $\gamma > 1$. The contribution of all recognized repetitions is the same if $\gamma = 1$. The interpretation of γ close to one is that the contribution of the past to the weight placed on each action is the same regardless of the time in the past the pattern appears. It means that the more frequently the pattern was followed by a certain outcome in the past, the stronger the belief in the occurrence of that outcome. For this reason, the parameter γ is called 'the Recency parameter'. It determines whether the individual is consistent with 'the law of primacy' ($\gamma > 1$) or 'the law of Recency' ($0 < \gamma < 1$).[3]

2.3 The decision rule

The decision rule for the two-parameter model is fairly simple, and uses only the payoff schedule as a source of 'external' influence on the human players. First, the beliefs of player i are calculated by normalizing the weight assigned to each pattern, $W^s_{i,t+1}$:

$$B^s_{i,t+1} = \frac{W^s_{i,t+1}}{\sum_S W^s_{i,t+1}} \tag{4}$$

Then, the expected payoffs are calculated, given these beliefs:

$$\hat{\pi}^s_{i,t+1} = E(\pi^s_{i,t+1}|B^s_{i,t+1}) = \sum_S (\pi^s_{i,t+1|x_{t+1}=s} B^s_{i,t+1}) \tag{5}$$

where $\pi^s_{i,t+1}$ is individual i's payoff from matching outcome s in time $t+1$. The expected payoffs are then normalized to form the probability assigned by individual i for each outcome:

$$P^s_{i,t+1} = \frac{\hat{\pi}^s_{i,t+1}}{\sum_S \hat{\pi}^s_{i,t+1}} \tag{6}$$

It should be noted, that the value of $\hat{\pi}^s_{i,t+1}$ by itself is not enough to predict individual i's choice. It is the value of $\hat{\pi}^s_{i,t+1}$ relative to all other values of $\hat{\pi}^{\overline{s}}_{i,t+1}$ (where $\overline{s}$ is all elements in S that are not s) that should be used for the prediction. We incorporate this issue by normalizing the expected payoffs (in equation 6).

3 Lund [18] was the first to discover the existence of these laws. See also [10,11,12,20,27].

3. The Four-Parameter Model

Both the two-parameter and four-parameter models have the same learning rule. They differ only in the decision rule. By adding two parameters, we create a decision rule that captures more than the influence of the payoff schedule. In fact, the proposed decision rule captures all other information that influences the human player, including influence on the subconscious, the kind of information that the player is not aware of or that seems irrelevant to her (her mood, for instance, or any other tendencies like one of the possible actions being the player's lucky number etc.).

The decision rule used by Cheung & Friedman (see [5], henceforth C&F) serves as an inspiration for the proposed decision rule in the four-parameter model. Their decision rule was based on the differences in expected payoffs between different actions, given the beliefs formed by the player. This difference in payoffs is then transformed using a cumulative distribution function (CDF). The CDF contains two parameters, α and β. The first controls for tendencies of the player towards (or away from) a certain action, and the second controls for the influence of the difference in payoffs between the actions.[4]

In our four-parameter model, the beliefs and expected payoffs are first calculated using the weights $W^s_{i,t+1}$, in the same way described in equations 4 and 5. Then, a function, $F(x)$, ("the mapping function") is introduced and utilized for a mapping process that converts the expected payoffs into probabilities that the individual assigns for the occurrence of each outcome. This mapping function, $F(x)$, should have the following characteristics:

$$\begin{aligned} & a \le x \le b \quad and \quad a < b \\ & \forall x_1, x_2 \in x; \quad x_1 > x_2 \quad \Rightarrow \quad F_{(x_1)} > F_{(x_2)} \\ & F_{(a)} = 0; \quad F_{(b)} = 1 \end{aligned} \tag{7}$$

The mapping is best described graphically in Fig. 1.

The expected payoffs are normalized [5] to fit into the domain of the mapping function. The probabilities assigned by the individual for the occurrence of each outcome appear on the range of the mapping function. A suitable mapping function is a function with parameters that have influence on the results (i.e., the probabilities assigned by the individual for the occurrence of each outcome). For our analysis, we utilized the Beta cumulative distribution function:

$$F(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} \int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt \tag{8}$$

where $B(\alpha,\beta)$ is the 'Beta function' with the following formula:

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt$$

4 Namely: $P(a_{it} = 1|\hat{r}_{it}; \alpha_i, \beta_i) = F(\alpha_i + \beta_i \hat{r}_{it})$ where $\hat{r}_{it}$ is the difference in expected payoffs. Note that this rule cannot be applied when at least one player faces more than two strategies. Furthermore, even if this setback could be somehow modified, the modified version of the rule involves using more than three parameters.

5 In case $b - a \neq \sum_S \hat{\pi}^s_{i,t+1}$ (where the domain of the mapping function is between a and b), then a normalization should be made to fit the expected payoffs into the domain of the mapping function. See Fig. 1 for clarifications. The normalization process will then be in the following way: $\tilde{\pi}^s_{i,t+1} = \frac{\hat{\pi}^s_{i,t+1}}{\sum_S \hat{\pi}^s_{i,t+1}}(b-a)$

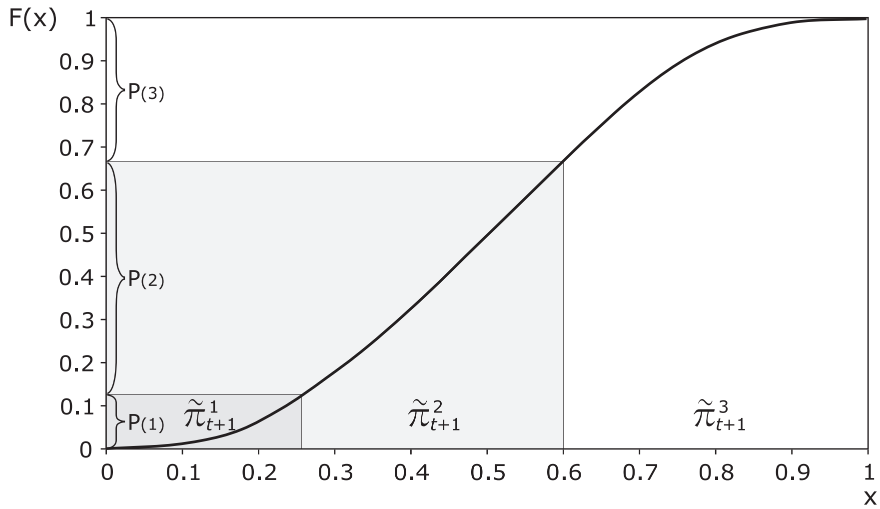

Fig. 1: An example of the mapping process of the two-parameter decision rule (applied in the four-parameter version) with $\alpha = \beta = 2$. The normalized expected payoffs are sorted on the X axis, and the Beta distribution function F(x) maps the expected payoffs into probabilities on the Y axis.

and α, β are the parameters of the model, as explained below. Note that the function satisfies the conditions set by equations 7. Now, it is possible to fit the expected payoffs to the range of the Beta CDF in order to map the probabilities for each choice. For that, the expected payoffs of individual i are normalized:

$$\tilde{\pi}^s_{i,t+1} = \frac{\hat{\pi}^s_{i,t+1}}{\sum\limits_S \hat{\pi}^s_{i,t+1}} \tag{9}$$

Now, the probability assigned by individual i of choosing each strategy can be calculated using $F(x)$ in the following way:

$$\begin{aligned} P^1_{i,t+1} &= F(\tilde{\pi}^1_{i,t+1}) \\ P^2_{i,t+1} &= F(\tilde{\pi}^1_{i,t+1} + \tilde{\pi}^2_{i,t+1}) - P^1_{i,t+1} \\ P^3_{i,t+1} &= 1 - \left[P^1_{i,t+1} + P^2_{i,t+1}\right] \end{aligned} \tag{10}$$

The structure of the Beta CDF and the normalization of $\hat{\pi}^s_{i,t+1}$ ensure the existence of:

$$\begin{aligned} &0 \le P_s \le 1 \quad \forall s \in S \\ &and \\ &\sum_S P_s = 1 \end{aligned} \tag{11}$$

The parameters α and β of the Beta CDF serve as parameters, and along with γ and δ form the four parameters of the model. The role of α and β is the same as in C&F. The change is in the structure. C&F use the parameter α to measure the tendency of the player to choose strategy 1, and the parameter β to measure the importance the player assigns to the differences between payoffs. Here, these elements will be determined by the mixture of the two parameters. For instance, when $\alpha = \beta = 1$, the probabilities will be determined only by the differences in payoffs, with the same ratio of probabilities as the ratio of differences, i.e. there is no influence of a decision rule in this case. When $\alpha = \beta = 1$, the Beta CDF represents a uniform distribution. The case of the uniform distribution is equivalent to $\alpha = 0, \beta = 1$ in C&F. In our model, a case of $\alpha = \beta >> 1$ is equivalent to a high α for the mid strategy (say, strategy 2 in the case of three strategies, or strategy 3 in the case of five). If this is the case, then the player has a tendency towards the strategy located in the middle of the strategy space. A case of $\alpha = \beta$ and α, β sufficiently close to zero is equivalent to high α's for the first and the last strategies. The greater the difference between α and β, the higher the strategies with the higher tendency will be located and vice versa. The strength assigned by the player to the differences in payoffs (the β in C&F) is also determined by the difference between the parameters α and β.

Fig. 1 illustrates the mapping process in a case of $S = 3$ and $\alpha = \beta = 2$. The normalized expected payoffs are ordered on the X axis, and the probability assigned to each action is the mapping of the normalized expected payoff on the Y axis using the Beta CDF function.

4. Experimental Design

The experiment was conducted at the computer laboratory of the economics department of Emory University. 45 subjects participated; all of them were undergraduate students from several disciplines. Each subject was randomly assigned to a computer and played a game. The game consisted of 200 rounds. In each round the computer generated a number, either 1, 2 or 3. Each participant was asked to guess the computer's outcome in each round before observing the outcome. For every correct match, the player received a payoff. No payoff was given for a wrong choice. The game lasted approximately 30 minutes, and the average payment was $12 per participant, including a $5 "show up" fee. The software used for this experiment was Excel, and all past choices of both players (the human and the computer) appeared on the screen and were available to the subjects for reference. The participants were informed that the computer was pre-programmed, and its choices were independent of the participant's choices (instructions provided to subjects are available from the author upon request). None of the players could influence the computer's outcomes.

Thirty out of 45 subjects were given a variation of a Markov series. In each round, the probability of facing each number (1, 2 or 3) was determined by the number appeared in the previous round. All 30 subjects received the same sequence. The probabilities of each number occurring are described in Table 1. These 30 subjects were assigned to two treatments, A and B. The two treatments differed by the payoff paid for a correct match.

Table 1: The Markov process that generates nature's decisions for treatments A and B.

Previous choice	P(1)	P(2)	P(3)
1	20%	30%	50%
2	50%	20%	30%
3	50%	30%	20%

Table 2 describes the payoff for each treatment. The remaining 15 subjects were assigned to treatment C. All subjects in treatment C faced a totally random process of generating outcomes (from this point, the word 'random' is used to describe the special case of a Markov series with equal probabilities for the occurrence of each action or choice). Their payoff schedule was the same as in treatment A.

Table 2: The payoff schedules of treatments A, B and C.

Treatment	Guessing 1 correctly	Guessing 2 correctly	Guessing 3 correctly
A,C	10¢	10¢	10¢
B	5¢	10¢	15¢

The difference between treatments A and B permits testing of the hypothesis that different payoff schedules influence the players' decision making. Since the only difference between treatment A and treatment B is the payoff schedule, one would expect to observe differences in parameters α and β among treatments. The learning parameters γ and δ, however, are expected to have the same values in both treatments. The reason for including treatment C is to have an additional, different environment of the input generating process. After all, one might claim that the existence of a Markov process is the reason the subjects made the decisions they did. A test in different environment can address such a claim.

5. Results

5.1 Overview of the data

Table 3 describes the number of times each subject chose each action. Table 4 describes the number of times all the players chose each action within blocks of 10 periods.

In both Table 3 and Table 4 the aggregate choices are presented in the last row for all players and for all rounds respectively. There is a difference in preference of choices between games, especially between game B and games A and C: action "1" was chosen more in games A and C (almost identical), where action "3" was chosen more in game B.

A chi-square test[6] was constructed to test the hypothesis that the distributions of choices aggregated over all participants within each treatment are the same across treatments for each block separately[7] (P values are available from the author upon request). After conducting three sets of tests (for differences between treatments A and B, A and C, and B and C separately), the findings for the whole 200 rounds are that the distributions of choices are different at the 5% significance level (which implies that subjects behaved differently in different treatments). The difference between treatments A and C is much lower, and the distributions of choices are not significantly different at the 1% significance level. The only logical explanation for the difference between treatment B and treatments A and C respectively is the difference in the payoff schedules. This means that the difference in the payoff schedule across treatments is the reason for the different behavior of subjects.

There is a difference in the distribution of choices between treatments A and B in 14 out of 20 blocks. The number of differences is even bigger between treatments B and C: 18 out of 20 blocks are significantly different. Finally, between treatments A and C, ten out of 20 blocks are significantly different.

6 See [24], p. 493 for description.

7 The alternative hypothesis is that the aggregate choices over all rounds in the two games are from different distributions.

Table 3: Frequencies of choices for every player in each game (aggregated over subjects). Note: for each treatment, a different group of subjects was used.

	Game A			Game B			Game C		
Player	**"1"**	**"2"**	**"3"**	**"1"**	**"2"**	**"3"**	**"1"**	**"2"**	**"3"**
1	72	74	54	60	65	75	77	68	55
2	54	83	63	26	25	149	53	102	45
3	69	65	66	32	84	84	71	71	58
4	60	80	60	53	68	79	62	75	63
5	101	44	55	31	56	113	64	69	67
6	115	43	42	56	70	74	69	73	58
7	62	81	57	70	61	69	60	71	69
8	67	70	63	40	43	117	70	66	64
9	62	63	75	4	72	124	58	76	66
10	73	64	63	52	70	78	55	69	76
11	0	0	200	67	62	71	71	64	65
12	56	80	64	38	58	104	63	76	61
13	75	66	59	25	96	79	50	49	101
14	65	54	81	26	71	103	58	74	68
15	51	97	52	1	10	189	65	69	66
Total	**982**	**964**	**1054**	**581**	**911**	**1508**	**946**	**1072**	**982**

Another interesting point worth mentioning is the location of the differences. The blocks that are significantly different between treatments A and B are located mainly in the last 100 rounds, whereas the blocks that are significantly different between treatments A and C are located mainly in the first 100 rounds. One can regard the first 100 rounds as a learning phase, where the matching process and the data are new to the subjects, and the last 100 rounds as the experienced phase. In this case, a difference in the first 100 rounds can be explained by some form of a learning stage, whereas a difference in the last 100 rounds can be seen as a real difference in decision making among experienced humans. According to these findings, one can conjecture that if there is an initial learning process, where individuals use the first few periods to get accustomed with nature's outcome rather than maximize payments, then this process is influenced by the observed outcomes and not by the payoff schedule.

We also conducted a chi-square test to test the hypotheses that individual choices are consistent with: (a) the Markov series, as generated by the empirical distribution of the computer's choices, (b) the actual frequencies of the computer's choices and (c) randomness. The results are presented in Table 5 (P values are available from the author upon request). Table 5 shows that the null hypothesis of consistency with the Markov series cannot be rejected for only 8 subjects from treatment C. 18 subjects were consistent with the actual frequencies chosen by the computer (3 subjects from treatment A, 2 from B and 13 from C) and 29 subjects were consistent with random decisions (10 subjects from treatment A, 6 subjects from B and 13 subjects from C). It can be seen from Table 5 that the decision making in treatment C is more consistent with all models than other treatments. Treatment C, being with equal payoffs for each action and equal probabilities for the occurrence of each action, encourage subjects to behave consistently with the random case. This explanation cannot be used for the consistency of the treatment with other models. We also tested the same

Table 4: Frequencies of choices aggregated among players for each game. A batch represents a block of 10 periods.

	Game A			Game B			Game C		
Block	**”1”**	**”2”**	**”3”**	**”1”**	**”2”**	**”3”**	**”1”**	**”2”**	**”3”**
1	40	54	56	36	65	49	63	39	48
2	74	34	42	54	30	66	37	52	61
3	47	45	58	31	48	71	39	68	43
4	34	40	76	26	40	84	54	49	47
5	38	60	52	38	56	56	58	56	36
6	48	53	49	33	46	71	53	58	39
7	41	49	60	33	53	64	49	60	41
8	55	42	53	31	49	70	21	53	76
9	43	47	60	29	29	92	60	50	40
10	34	63	53	24	54	72	51	51	48
11	42	48	60	24	53	73	53	38	59
12	46	57	47	30	46	74	49	54	47
13	47	59	44	22	54	74	49	52	49
14	59	51	40	22	42	86	53	47	50
15	59	38	53	26	39	85	49	49	52
16	56	53	41	17	48	85	21	62	67
17	52	42	56	20	49	81	52	65	33
18	66	48	36	26	40	84	46	57	47
19	51	36	63	32	43	75	40	71	39
20	50	45	55	27	27	96	49	41	60
Total	**982**	**964**	**1054**	**581**	**911**	**1508**	**946**	**1072**	**982**

hypotheses by using the last 100 rounds for each subject, to account for any learning process. Table 5 shows that the null hypothesis of consistency with the Markov process cannot be rejected for 11 subjects (10 subjects from treatment C, and one subject from A). We find even fewer subjects that are consistent with either random decision or the actual frequencies.

Table 5: The number of subjects consistent with realizing the actual Markov series (”Markov”), the actual frequencies (”Actual freq.”) and random choices (”Random”) for all rounds (200) and the last 100 rounds (100).

	Markov (100)	**Markov (200)**	**Actual freq. (100)**	**Actual freq. (200)**	**Random (100)**	**Random (200)**
Treatment A	1	0	0	3	9	10
Treatment B	0	0	0	2	5	6
Treatment C	10	8	11	13	11	13

In summary, there is a difference in decisions among players from different treatments. One reason is the payoff schedule, but there is also a difference as a result of nature’s choices.

5.2 Procedure

We first estimate the values of the parameters in our model for each subject separately. For convenience, we restrict all the parameters to values higher than zero and smaller than five[8]. This restriction ensures the necessity that the decision parameters should be strictly positive. The parameters are estimated for each player i separately by minimizing the mean of absolute distances between the actual probability distribution and the one generated by the model in each round:

$$[\alpha_i^*, \beta_i^*, \gamma_i^*, \delta_i^*] = \arg\min_{\alpha,\beta,\gamma,\delta} \left\{ \frac{1}{T} \sum_t \left| f_i^{observed} - f_i^{predicted} \right| \right\} \tag{12}$$

where the observed probability distribution for each player in each round is such that the probability of choosing the chosen action by the player is 1, and 0 for the other actions. We also test in the same way the fictitious play model [1] (as presented in [17], page 31), the hypothesis that players recognize the actual Markov process generated by nature's actual choices, the hypothesis that players choose actions randomly with equal probability in each action and the Erev & Roth one parameter model. The model based on the Cournot dynamics [7] and the model of Sarin & Vahid [23], being deterministic models and not probabilistic, are not tested here.

It should be mentioned that when minimizing the distances for each period separately, the expected probability distribution is generated by the model, while the observed probability distribution is unknown. All that can be observed is a distribution with the value "1" for the chosen action, and "0" for all other actions. By minimizing the mean of absolute distances instead of the squared distances, only the chosen action is being tested, since there is no information regarding actions that were not chosen. For this reason, we minimized the mean absolute distances and not the mean square distances.

5.3 Parameter estimation and comparison

Table 6, Table 7 and Table 8 present the scores obtained by the two versions of the pattern recognition model and the other models mentioned above. P-R(2) is the two-parameter pattern recognition model and P-R(4) is the four-parameter model. F-P stands for Brown's fictitious play, 'Markov' represents the case in which players identify the actual Markov process that generates the computer's outcomes, 'Random' is the assumption that players pick actions at random and E-R is Erev & Roth's one parameter reinforcement model. From Table 6, Table 7 and Table 8 it can be seen that the four-parameter version of the pattern recognition model performs better than other models for 44 players (98%), and the two-parameter version performs better than others for 36 players (80%).

Since the scoring rule we are using is the average of all scores over periods for each player, a t-test is conducted to test the hypothesis that for each player, the two versions of the pattern recognition model perform significantly better than the other competing models

8 The upper bound is meant to shorten the data processing time. The difference between a value higher than five and a value of five create differences that can be easily ignored, on the basis that a value of five or higher has the same intuition. For example, the two cases $\alpha = 5$ and $\alpha = 7$ both imply a strong recency effect. This implication by itself is more important than the actual value of γ. Regarding the learning parameters, there is no reason to believe that no finite value will be found. If this was the case then the model converges to random choice. Since all scores of the model are lower than the scores of the random choice (Tables 6, 7 and 8), a finite value of the learning parameters exists.

Table 6: Mean absolute distance deviation for each player. Treatment A

Player	Random (equiprobable)	Markov	F-P	E-R	P-R(2)	P-R(4)
1	0.4444	0.4292	0.4404	0.4444	0.3972	0.3941
2	0.4444	0.4436	0.4486	0.4442	0.4126	0.4045
3	0.4444	0.4328	0.4473	0.4445	0.4010	0.3939
4	0.4444	0.4475	0.4504	0.4446	0.4205	0.4178
5	0.4444	0.4324	0.4410	0.4160	0.3816	0.3301
6	0.4444	0.4309	0.4306	0.3963	0.3661	0.2760
7	0.4444	0.4394	0.4426	0.4379	0.4248	0.4115
8	0.4444	0.4518	0.4454	0.4445	0.4369	0.4276
9	0.4444	0.4372	0.4414	0.4380	0.4194	0.4128
10	0.4444	0.4346	0.4393	0.4423	0.4038	0.4009
11	0.4444	0.4540	0.4636	0.0089	0.4359	0.0233
12	0.4444	0.4399	0.4485	0.4445	0.4179	0.4033
13	0.4444	0.4222	0.4433	0.4445	0.3777	0.3675
14	0.4444	0.4446	0.4491	0.4402	0.4194	0.3698
15	0.4444	0.4352	0.4492	0.4224	0.4084	0.3978

Table 7: Mean absolute distance deviation for each player. Treatment B

Player	Random (equiprobable)	Markov	F-P	E-R	P-R(2)	P-R(4)
16	0.4444	0.4527	0.4449	0.4386	0.4173	0.3848
17	0.4444	0.4395	0.4528	0.2775	0.3563	0.1611
18	0.4444	0.4417	0.4546	0.4171	0.3868	0.3596
19	0.4444	0.4219	0.4474	0.4363	0.3884	0.3722
20	0.4444	0.4351	0.4574	0.3547	0.3901	0.2870
21	0.4444	0.4351	0.4474	0.4351	0.4077	0.3887
22	0.4444	0.4337	0.4416	0.4400	0.4070	0.3879
23	0.4444	0.4269	0.4430	0.3529	0.3460	0.2556
24	0.4444	0.4537	0.4621	0.3274	0.3892	0.2619
25	0.4444	0.4518	0.4470	0.4304	0.4287	0.3844
26	0.4444	0.4408	0.4448	0.4435	0.3846	0.3762
27	0.4444	0.4393	0.4581	0.3782	0.4006	0.3107
28	0.4444	0.4540	0.4512	0.3957	0.4105	0.3703
29	0.4444	0.4476	0.4519	0.3939	0.4018	0.3097
30	0.4444	0.4526	0.4608	0.0614	0.3397	0.0551

(at the 5% significance level). In other words, we tested whether the lower score of the pattern recognition model relative to each of the other models represents a significantly lower average.

Table 8: Mean absolute distance deviation for each player. Treatment C

Player	Random (equiprobable)	Markov	F-P	E-R	P-R(2)	P-R(4)
31	0.4444	0.4279	0.4445	0.4440	0.3944	0.3719
32	0.4444	0.4224	0.4414	0.3815	0.3682	0.3490
33	0.4444	0.4354	0.4448	0.4444	0.3811	0.3647
34	0.4444	0.4246	0.4500	0.4444	0.3447	0.3240
35	0.4444	0.4232	0.4456	0.4444	0.3929	0.3826
36	0.4444	0.4237	0.4422	0.4444	0.3630	0.3574
37	0.4444	0.4286	0.4456	0.4444	0.3592	0.3540
38	0.4444	0.4213	0.4443	0.4444	0.3110	0.3020
39	0.4444	0.4366	0.4499	0.4444	0.3946	0.3869
40	0.4444	0.4267	0.4470	0.4444	0.3907	0.3738
41	0.4444	0.4164	0.4419	0.4433	0.3490	0.3296
42	0.4444	0.4320	0.4467	0.4426	0.3644	0.3575
43	0.4444	0.4250	0.4383	0.4246	0.3725	0.3084
44	0.4444	0.4124	0.4462	0.4438	0.3335	0.3282
45	0.4444	0.4369	0.4405	0.4435	0.3618	0.3377

Table 9 shows the results. Comparing the two-parameter and the four-parameter versions of the pattern recognition model to E-R, the performance of the pattern recognition model is significantly better for 22 and 37 players respectively (out of 45 players). The relative performance of the pattern recognition model is even better when compared with the other models.

Table 9: Number of players who preformed significantly better (on average over periods) in the pattern recognition model, compared to the 'random', 'Markov', 'F-P' and 'ER' respectively

	Random	Markov	F-P	E-R
P-R(2)	34	36	37	22
P-R(4)	40	39	40	37

Fig. 2 and Fig. 3 describe the distribution of the learning parameters γ and δ in histograms for the two-parameter and the four-parameter models respectively. Each histogram represents the distribution of each parameter in each treatment (parameters values are available from the author).

In the two-parameter model, the distributions of δ are similar in all three treatments. A t-test is used to test the hypothesis that the mean value of each parameter is equal across treatments. The null hypothesis for all comparisons cannot be rejected (P-values are higher than 0.50). There are high concentrations close to zero and in high values. This suggests that there are two types of individuals: those who look for patterns, and those who do not care much about them (or they do care, but do not succeed finding them).

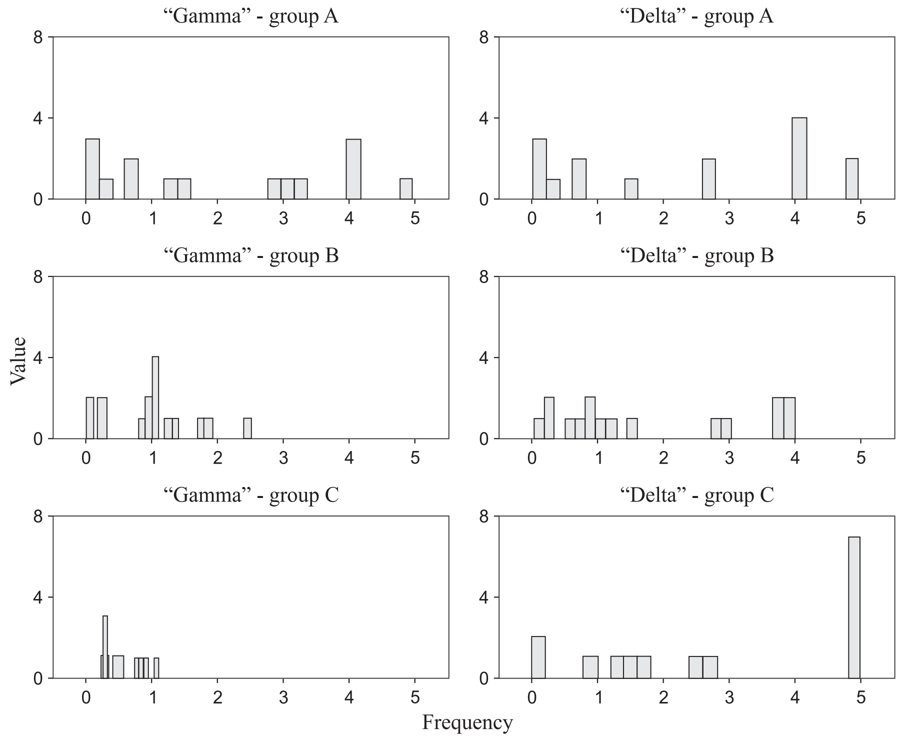

Fig. 2: The histograms of the learning parameters Gamma and Delta for each treatment. The four-parameter model.

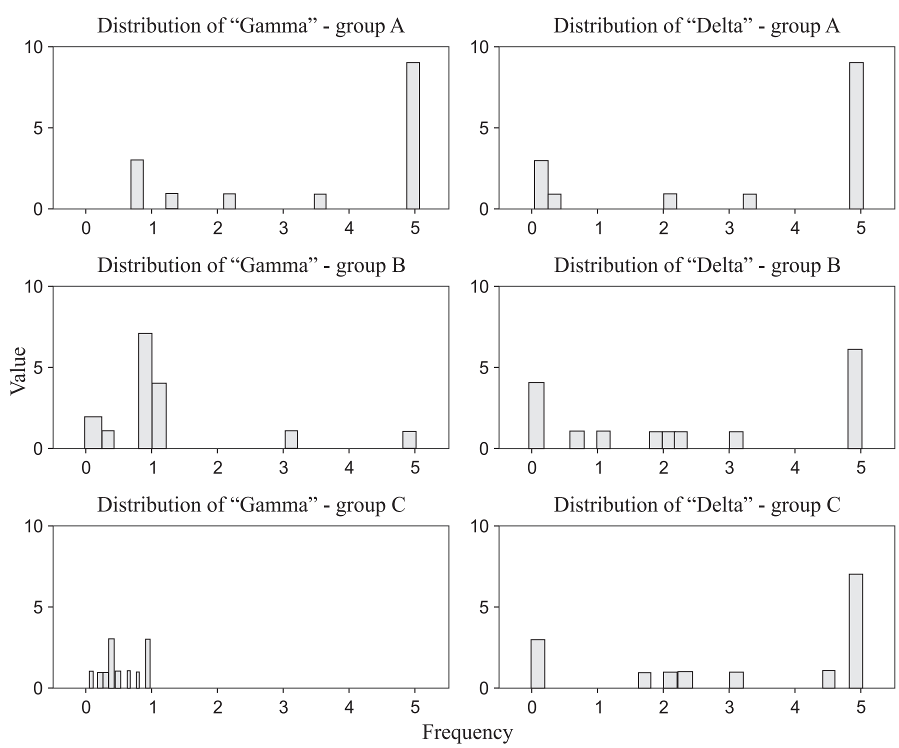

Fig. 3: The histograms of the learning parameters Gamma and Delta for each treatment. The two-parameter model.

The distributions of γ, however, are not that similar. From Fig. 2, the mean is the highest in treatment A (3.61), and the lowest in treatment C (0.53). In treatment B, the mean is 0.94. The t-test rejects the null hypothesis of equality of means across treatments. The reason might be the incentive to look into the past. In treatment C, where the choices are random, there are fewer repeated patterns to be found, because of the random process. Therefore, an individual who recognizes patterns finds it unnecessary to waste time and energy on searching back in the past. This is why subjects are consistent with the recency effect ($0 < \gamma < 1$). In treatment A, on the other hand, the Markov process generates a higher number of repeated patterns with length $n_j > 2$. Individuals who recognize these repeated patterns tend to look back and use this information. This is the reason for the consistency of subjects in this treatment with the primacy effect ($\gamma > 1$). Although the pattern generated by the computer in treatment A is similar to the pattern in B, players are influenced to choose action "3" more in treatment B because of the different payoff schedule. They have more incentive to choose action "3", which can drive them away from the search for patterns. This 'neutrality' is consistent with $\gamma \cong 1$.

In the four-parameter model the story is quite the same for the distributions of γ, but not for the distributions of δ. The distributions of δ in the four-parameter version have higher averages than in the two-parameter version in all treatments, although the variance is almost the same. Using the same t-test, the hypothesis that the mean is the same cannot be rejected (P-values are higher than 0.05), although high concentrations of relatively high and relatively low values cannot be found. The distributions of γ have the same properties as in the two-parameter model. The mean is the highest in treatment A (2.65), lower in B (1.18) and the lowest in C (0.56).

We define λ_i^s as the tendencies of player i for each action s. A positive value means that player i has a tendency towards action s. A negative value implies a tendency away from action s. The tendencies of player i to (or away from) action s is calculated as follows:

$$\begin{aligned} \lambda_i^1 &= F_{(\frac{1}{3}|\alpha,\beta)} - \frac{1}{3} \\ \lambda_i^2 &= F_{(\frac{2}{3}|\alpha,\beta)} - F_{(\frac{1}{3}|\alpha,\beta)} - \frac{1}{3} \\ \lambda_i^3 &= 1 - F_{(\frac{2}{3}|\alpha,\beta)} - \frac{1}{3} = \frac{2}{3} - F_{(\frac{2}{3}|\alpha,\beta)} \end{aligned} \tag{13}$$

Fig. 4 and Fig. 5 present the distributions of tendencies among players in the four-parameter model. To calculate the tendencies, the decision parameters α and β for each player were used (values are available from the author) and constructed the Beta cumulative distribution function $F_{(x)}$.

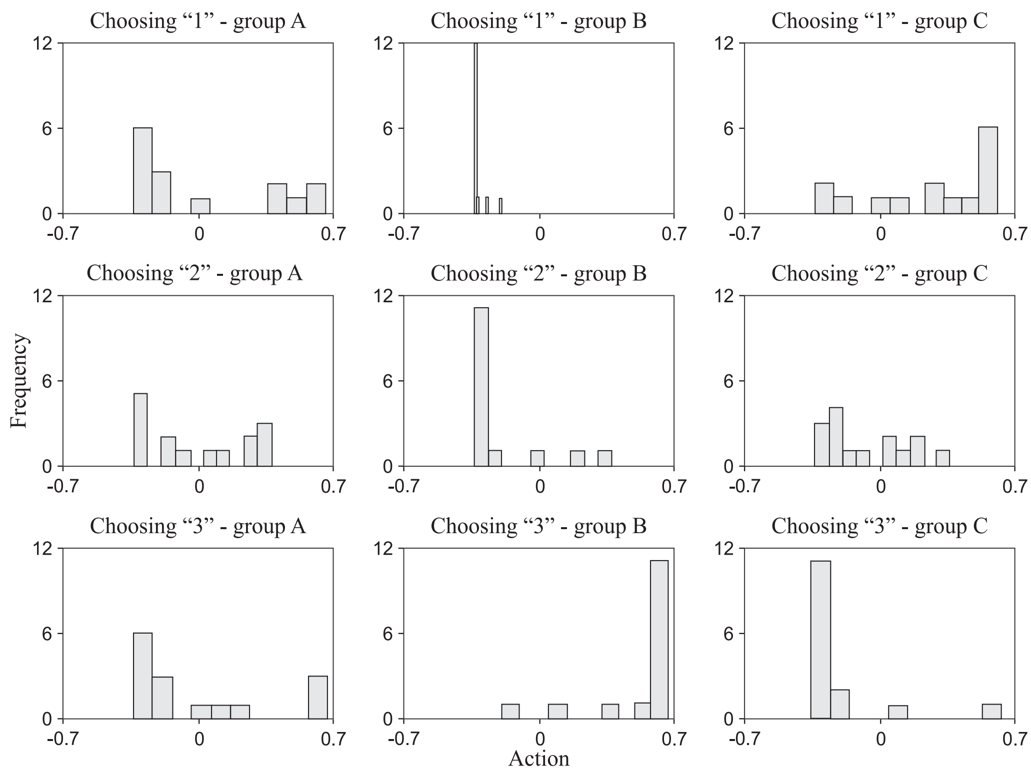

Fig. 4: Histograms of level of tendencies towards (or away from) actions for each action in each treatment.

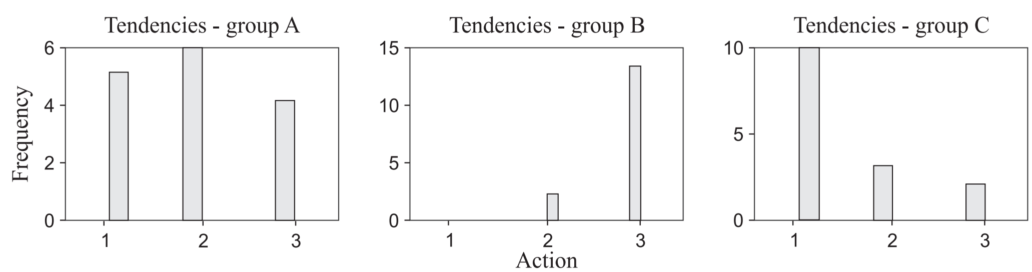

Fig. 5: The distributions of actual tendencies, among players in each treatment.

A simple t-test was used to test the hypothesis that the tendency distributions for each action in each treatment have zero mean. The results are shown in Table 10, where the level of significance is $\alpha = 0.05$. As expected, the null hypothesis for all distributions cannot be rejected in treatment A. This means that on average, players in treatment A do not have tendencies towards (or away from) any action. These results are consistent with the symmetric payoff schedule faced by the player in treatment A. Treatment B, on the other hand, has strong tendencies towards action ”3”, and away from actions ”1” and ”2”. the mean tendency for actions ”1” and ”2” are significantly negative (which means that on average, players’ tendencies in treatment B are away from action ”1” and ”2”), while the mean tendency for action ”3” is significantly positive (tendency towards action ”3”). These results are also consistent with the payoff schedule faced by treatment B, where the payoff for guessing action ”3” is the highest, and that for guessing action ”1” is the lowest. The results from treatment C are surprising. Although one would expect tendencies in this

Table 10: P-values for the null hypothesis that the tendencies destributions for each action and in each treatment has a zero mean.

	Action ”1”	**Action ”2”**	**Action ”3”**
Treatment A	0.5126	0.4148	0.8806
Treatment B	0.0000	0.0009	0.0000
Treatment C	0.0069	0.2329	0.0110

treatment to have the same mean as in treatment A (because the payoff schedules are the same), there is a tendency towards action ”1” and away from action ”3”.

6. Conclusions

The process of pattern recognition is well known to science, and is a focus of research in fields such as mathematics, computer science, biology and psychology. There has been limited theoretical or empirical investigation of pattern recognition in economics to date, despite its obvious existence in individual decisions and aggregate variables. The goal of this paper was to present a model that provides an empirical fit to the actual incentivized behavior of individuals by allowing for pattern recognition.

We present a learning model that captures pattern recognition as a feature of a process of learning. In this regard it differs from other learning models, in that it explicitly uses historical dynamics in decision making, and therefore employs all past information, including the sequence of prior outcomes, in the learning process. Individuals using the pattern recognition learning rule observe all of the past, but decide what segments of history to use, and what not to use. This choice is dynamic in the sense that it is updated after every period. After scanning all the information they have, they sample the parts that look essential (because they are consistent with the most recent sequence of activity), and abandon all the rest until the next period.

The learning rule is accompanied by a decision rule that takes into account exogenous biases of individuals towards (or away from) certain actions. This decision rule also captures the influence of the payoff schedule on decisions. This rule has the property that the number of parameters determining an individual's decision is constant, regardless of the number of possible actions nature generates. Thus, it keeps the total number of parameters in the whole model constant (for each of the two versions of the model), no matter how large the set of actions is for the individual.

The results show that the pattern recognition model provides a better account of decisions than several well-known existing learning models, for the situation studied (this is true even when there is no actual, deterministic, pattern to realize). It embodies a fairly intuitive behavioral concept that seems appropriate to the particular simple and generic environment studied (as well as others, potentially). Interesting research can be conducted in testing this model on players' behavior in games where they face human opponents rather than nature.

References

[1] Brown, G. (1951), ”Iterated Solution of Games by Fictitious Play”, in ”Activity Analyses of Production and Allocation (T. Koopmans, Ed), pp. 377-380. New York: Wiley.
[2] Brown, Gregory W. and Michael T. Cliff (2005), ”Investor Sentiment and Asset Valuation,” Journal of Business, 78(2), 405-440.
[3] Camerer, Colin and Ho, Teck-Hua (1999), ”Experienced-Weighted Attraction Learning in Normal Form Games”, Econometrica, 67(4), 827-874.
[4] Cheng, B and Titterington, D.M (1994), ”Neural Networks: a Review” from Statistical Perspective, Statistical Science, 9(1), 2-54.
[5] Cheung, Yin-Wong and Friedman, Daniel (1997), ”Individual Learning in Normal Form Games: Some Laboratory Results”, Games and Economic Behavior, 19, 46-76.
[6] Cosman, P.C, Oehler, K.L, Riskin, E.a and Gray, R.M (1993), ”Using Vector quantization for Image Processing” Proc. IEEE, 81, 1,326-1,341.
[7] Cournot, A. (1838), ”Researches into the Mathematical Principles of the Theory of Wealth” (English Ed), New york: Macmillan, 1897.
[8] Egmont-Peterson, M, De Ridde, D and Handels, H (2002), ”Image Processing Using Neural Networks - a Review” Pattern Recognition, 35(10), 2,279-2,301
[9] Erev, Ido and Roth, Alvin, E. (1998), ”Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Equilibrea”. The American Economic Review, 88(4), 848-881.
[10] Fornham, Adrian (1986), ”The Robustness of the Recency Effect: Studies Using Legal Evidence”, The Journal of General Psychology, 113(4), 351-357.
[11] Hovland, Carl I, Mandell, Wallace (1957), ”Is There a Law of Primacy in Persuasion?”, The Order of Presentation in Persuasion, Crl I. Hovland et al. New Haven: Yale University Press.
[12] Hovland, Carl I, Mandell, Wallace, Campbelll, Enid H, Brock, Timothy, Luchins, Abraham S, Cohen, Arthur R, McGuire, William J, Janis, Irvin L, Feierabend, Rosalind A and Anderson, Norman H (1957), ”The Order of Presentation in Persuasion”, New Haven: Yale University Press.
[13] Huck, Steffen, Normann, Hans-Theo and Oechssler, Jörg (1999), Learning in Cournot Oligopoly - an Experiment”, The Economic Journal, 109, C80-C95.
[14] Jain, A.k, Duin, U.P.W and Mao, J (2000), ”Pattern Recognition: a Review”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1), 4-37.
[15] Lee, Charles, Andrei Shleifer and Richard Thaler (1991), ”Investor Sentiment and the Closed-End Fund Puzzle,” Journal of Finance, 46, 75-109.
[16] Lee, W. Y., Jiang, C. X. and Indro, D.C. (2002), ”Stock Market Volatility, Excess Returns and the Role of Investor Sentiment,” Journal of Banking and Finance, 26 (12), 2277-2299.
[17] Levine, David K. and Fudenberg, Drew (1998), ”The Theory of Learning in Games”, The MIT Press, Cambridge, MA, London, England.
[18] Lund, Frederick H. (1925) ”The Psychology of Belief: a Study of Its Emotional and Volitional Determinants”, Journal of Abnormal and Social Psychology, 20, 174-196.
[19] Margolis, Howard (1987),”Patterns, Thinking and Cognition, a Theory of Judgment”, the University of Chicago Press, Chicago and London.
[20] Miller, Norman and Campbell, Donald T. (1959), ”Recency and Primacy in Persuasion as a Funcion of the Timing of Speeches and Measurements”, Journal of Abnormal and Social Psychology, 59(1), 1-9.
[21] Nagel, Rosemerie (1995), Unraveling in Guessing Games: An Experimental Study”, The American Economic Review, 85(5), 1313-1326.
[22] Rötheli, Tobias F. (1998), ”Pattern Recognition and Procedurally Rational Expectations”, Journal of Economic Behavior & Organization, 37, 71-90.
[23] Sarin, Rajiv and Vahid, Farshid (1999), ”Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice”, Games and Economic Behavior, 28, 294-309.
[24] Sheskin, David J. (2004), ”Handbook of Parametric and Nonparpmetric Statictical Procedures”, Third Edition, Chapman & Hall/CRC
[25] Sonsino, Doron (1997), ”Learning to Learn, Pattern Recognition and Nash Equilibrium”, Games and Economic Behavior, 18, 286-331.
[26] Stahl, Dale O. (1996), ”Boundedly Rational Rule Learning in a Guessing Game”, Games and Economic Behavior, 16, 303-330.
[27] Weld, H. and Roff, M. (1938), A Study in the Formation of Opinion Based Upon Legal Evidence, American Journal of Psychology, 51, 609-628.