ISSN 1440-771X
ISBN 0 7326 1088 5

Strategy Similarity and Coordination

Rajiv Sarin and Farshid Vahid

Working Paper 8/2001

2001

DEPARTMENT OF ECONOMETRICS AND BUSINESS STATISTICS

MONASH UNIVERSITY

AUSTRALIA

Strategy Similarity and Coordination[1]

Rajiv Sarin
Department of Economics
Texas A&M University
College Station, TX 77843
USA

Farshid Vahid
Department of Econometrics
Monash University
Clayton, Victoria 3168
Australia

July 2001

[1] This paper was formerly titled, "Payoff Assessments without Probabilities: Incorporating "Similarity" among Strategies."

Abstract

This paper introduces similarity among strategies in the payoff assessment model of choice (Sarin and Vahid (1999, *GEB*)). The assessments of strategies that are more similar to the chosen strategy are updated more similarly to the chosen strategy. We use this model to explain a recent experiment. The coordination game repeatedly played by the experimental subjects had two symmetric, efficient and strict stage game Nash equilibria. In the experiment, the subjects always converged to play one of these equilibria, and converged to this equilibrium remarkably fast. The model we propose converges to choose the same equilibrium, and does so in roughly the same number of repetitions. Statistical tests are performed to distinguish between the choice distributions generated by the model and the observed choice distributions.

1 Introduction

In many economic environments decision makers have very little information about the environment in which they have to make decisions and obtain very little information with each repetition of the environment. Examples include firms choosing between prices, quantities or technologies, and consumers choosing between car mechanics or restaurants. Among these decision problems there are many in which the choices among different strategies can be ordered according to how (subjectively) similar one strategy is to another. For a firm, choosing output levels that are close together would appear to be a more similar strategy than choosing output levels that are not close together. A consumer choosing among restaurants may view a Chinese restaurant as being more similar to a Thai restaurant than to a French restaurant.

Decision makers may use similarity among strategies to simplify the problem they face, and the manner in which they simplify the problem may have consequences for the equilibrium they "learn" to play. In this paper we propose a model of the process by which players learn to play the game in which they may know very little about the game, and observe very little with each repetition, and in which strategy labels provide some information about the similarity among the strategies. We use this model to understand an experiment conducted by Van Huyck, Battalio and Rankin (1996).

The model we develop combines features of two recent choice theoretic models: The payoff assessment model (Sarin and Vahid (1999)) and the model of case-based decision theory (Gilboa and Schmeidler (1995, 1997)). In accordance with the former we suppose that the decision maker associates with each of her strategies a subjectively assessed payoff, which represents the payoff the players "expects" or "anticipates" she will receive from the choice of that strategy. At each stage, the decision maker chooses the strategy she assesses to have the highest payoff.

After the agent chooses a strategy, she obtains a payoff. This is the objective payoff to the decision maker. She uses this objective payoff to update her earlier assessments. Whereas in Sarin and Vahid (1999) the agent was assumed to only update her assessments of the chosen strategy,[1] we shall suppose the agent also updates her assessments of strategies she sees as "similar" to the chosen strategy. Our use of the idea of "similarity" between strategies was inspired by the work of Gilboa and Schmeidler. This

[1] Implicitly, there was no similarity between actions.

idea has been used earlier by Rubinstein (1988), who uses it to explain choice behavior among lotteries. As in Rubinstein we also suppose that the agent uses similarity among strategies to simplify the decision problem she faces. In contrast to Rubinstein, we allow for different degrees of similarity among the strategies. The assessments of strategies with greater similarity being updated in a more similar manner.

In the Van Huyck, Battalio and Rankin (henceforth VBR) experiment, groups of people repeatedly played a large coordination game for 75 periods. The game was played by 5 players, each of whom had 101 strategies. The subjects did not know the payoff function of the game they were playing, or even the strategies available to the other players. At each repetition of the game each subject observed only the payoff she obtained from the strategy chosen. Each of the strategies of the players was labelled in the same way across the players. Specifically, the 101 strategies for each player were labelled 0, 1, ..., 99, 100. This labelling of the strategies leads to a very natural similarity notion across strategies: The strategy 8 would appear to be closer to 7 than strategy 30.[2] The payoff function incorporated the same kind of similarity among strategies.

The stage game the subjects played had two strict Nash equilibria. The equilibria were symmetric and efficient. Hence, traditional refinements could not reduce the set of the Nash equilibria. The experimental subjects, however, managed to coordinate in each experimental session on the same equilibrium. The subjects were able to consistently choose among the equilibria, even though they had such little information about the game and observed so little with each repetition of the stage game. Moreover, the subjects managed to coordinate on the equilibrium in a remarkably short time: Typically the median player had chosen the equilibrium strategy in the first 25 periods and had converged to choose only this strategy within 40 periods.

VBR attempted to explain these remarkable results by looking at the selection among equilibria implied by alternate adaptive selection dynamics. VBR showed that dynamic learning processes, like myopic best response dynamics, provided little help in discriminating between the equilibria. Their simulations of a model of reinforcement learning due to Cross (1973), revealed that it could generate a final distribution of choices similar to those of the experimental subjects. However, it took the Cross model no less than 750 periods to exhibit such a final distribution.

[2] We think of this notion of closeness as natural because we grow up using natural numbers (in playing card games, counting sheep and money, etc.) and they have this notion of closeness.

In contrast to the VBR analysis, we obtain striking support for the model we propose in this paper. The median player converges to choose the same equilibrium as chosen by the experimental subjects. Furthermore, the model converges to that equilibrium in *roughly the same* number of repetitions as did the experimental subjects. Statistical tests cannot distinguish the choice distributions generated by the model and the observed choice distributions in *almost every period*. Our findings suggest that people do use similarity among strategies to simplify large strategy spaces when there is incomplete information about the payoff function. Furthermore, our findings suggest that "close" similarity functions may have fairly different performance.

Given the success of introducing similarity among strategies we attempted to modify the Cross model to include it also. We also modified the Cross model to allow the step-size or learning rate to decrease over time. With the second modification the Cross learning model becomes very much like the basic model studied in Erev and Roth (1998), even while it has more attractive short-run properties.[3] Surprisingly, neither modification improves the predictive power of the Cross model.

The updating of unplayed, but related, strategies has also recently been found useful in describing data in Cournot oligopoly experiments. Huck, Norman and Oechssler (1999) show that Cournot oligopolists with very limited information of the payoff function may also update the attractiveness of unplayed related strategies. Specifically, they find that if the choice of a higher strategy (quantity) results in higher payoffs (profits), then the agent behaves as if (s)he has updated (increased) the payoff assessment of other higher strategies.

This paper is structured as follows. The next section presents the model of this paper. Section 3 summarizes the VBR experiments. In Section 4 we present simulations of our model and compare the results with those obtained in the VBR experiments. Also, in Section 4, we estimate the parameters of our model which lead it to fit best with the entire path of play in the experimental sessions. Section 5 considers variants of the Cross model and compares their predictions with the experimental data. Section 6 concludes.

[3]These are largely a consequence of it being linear in payoffs. See Borgers, Morales and Sarin (2001) for the details.

2 The Model

Let I denote the set of players in the game, and let S^i denote the pure strategy set of player $i \in I$. We shall assume that there are a finite number m of players. For notational simplicity, suppose that each player has a finite number J of strategies. By s we shall denote the strategy profile $\left(s^1, ..., s^m\right)$ actually played by the m players. Denote by $u^i_j\left(n\right)$ the subjective assessment of player i for her strategy j, s^i_j, at repetition n, and let $u^i\left(n\right)$ denote the vector of her subjective payoff assessments of the payoff for each of her strategies. Hence, at time n, the game player i subjectively assesses that she is playing is given by $\Gamma^i\left(n\right) = \left(S^i,\, u^i\left(n\right)\right)$.

At each n, each player i chooses the strategy that has the highest assessment. Let $s^i_{\max}\left(n\right)$ denote the strategy that player i assesses to have the highest payoff at time n. After each player i at time n chooses $s^i_{\max}\left(n\right)$, she obtains a payoff $\pi^i\left(s^i_{\max}\left(n\right),\, s^{-i}_{\max}\left(n\right)\right)$. After receiving the payoff each player updates her assessments. We suppose that the individual takes a weighted average of her assessment and the payoff she receives to form her updated assessment. Specifically, if at time n player i chooses strategy $s_j = s^i_{\max}\left(n\right)$, and receives a payoff π^i_j, then

$$u^i_j\left(n+1\right) = u^i_j\left(n\right) + \lambda\left(\pi^i_j - u^i_j\left(n\right)\right) \qquad 0 < \lambda < 1 \tag{1}$$

We now discuss how player i updates the assessments of her unplayed strategies when she chooses s^i_j and receives π^i_j at time n, and that both strategies s^i_k and s^i_l are not played at round n. If the decision maker thinks that s^i_k is more similar to s^i_j than she thinks s^i_l is to s^i_j then we want that she update her assessment of s^i_k in a manner more similar to how she updates the assessment of s^i_j than how she updates the assessment of s^i_l. That is, the player updates her assessments of strategies which she thinks are more similar in a more similar way.

Formally, for each player i we consider a similarity function between strategies, $f^i : S^i \times S^i \rightarrow \Re$. A value of f^i of 1 indicating that the player sees the two strategies as identical, and a value of 0 indicating that she sees no similarity between them at all. We may now write the manner in which the agent updates all her assessments. If at time n player i chooses strategy s^i_j and receives a payoff of π^i_j, then:

$$u^i_k\left(n+1\right) = u^i_k\left(n\right) + \lambda f^i\left(s^i_j,\, s^i_k\right)\left(\pi^i_j - u^i_k\left(n\right)\right) \qquad \forall s^i_k \in S^i \tag{2}$$

Note this (2) collapses to (1) if there is no similarity among actions, i.e. $f^i\left(s_j^i, s_k^i\right) = 0$ for $s_j^i \neq s_k^i$, because a strategy is identical to itself, i.e. $f^i\left(s_j^i, s_j^i\right) = 1$.

The model implies that the assessment of a played action is moved in the direction of the obtained payoff by a fraction λ. The assessments of unplayed actions also are updated. The payoff received from the played action, which provides the only new source of information, is used to update the assessments of unplayed actions by an amount λf.

Observe that we assume that the similarity function is constant over time. This may appear to be a restrictive assumption if the strategy space is small, the environment is stationary, and the decision maker knows that she will be facing the same decision problem sufficiently often to experience each of the available strategies many times. However, for an agent lacking in any specific knowledge about the stationarity of the environment and facing a large set of strategies, it is intuitive to consider, for example, that the strategy labelled 19 stays close to strategy labelled 20 no matter how non-stationary the environment is over time. In such situations it does not appear overly restrictive to suppose that the similarity function is constant over time. Such an assumption on the similarity function also greatly simplifies the analysis and facilitates its estimation.

We consider two types of similarity functions in this paper. Both assume that the strategies can be represented by numbers. In both any player i conjectures that all strategies within a symmetric "window" around the played strategy s_j^i to be similar to s_k^i, and that more distant strategies are less similar. The size of the window is the single parameter that describes these similarity functions. The first similarity function we consider, which we call the Bartlett similarity function,[4] supposes that the degree of similarity decreases linearly as the distance between the played strategy s_j^i and others increases. In the second similarity function we consider, which we call the Parzen similarity function[5] similarity decreases less than linearly for closer strategies and more than linearly for farther ones.

[4] This similarity function is inspired by the "Bartlett window" (see, e.g., Brockwell and Davis (1991), p.361-2).

[5] This similarity function is inspired by the Parzen window (see, e.g., Brockwell and Davis (1991), p.361-2).

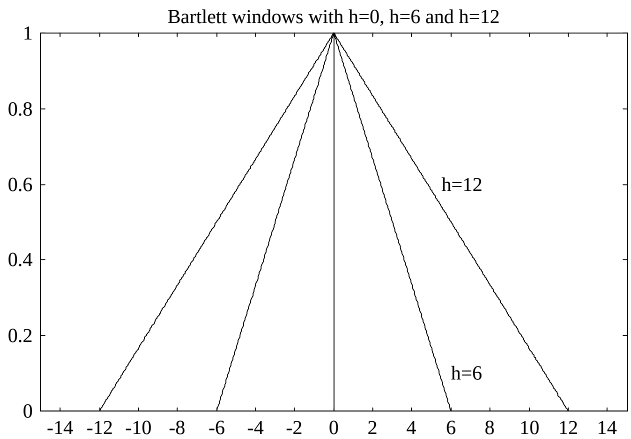
The Bartlett similarity function is given by:

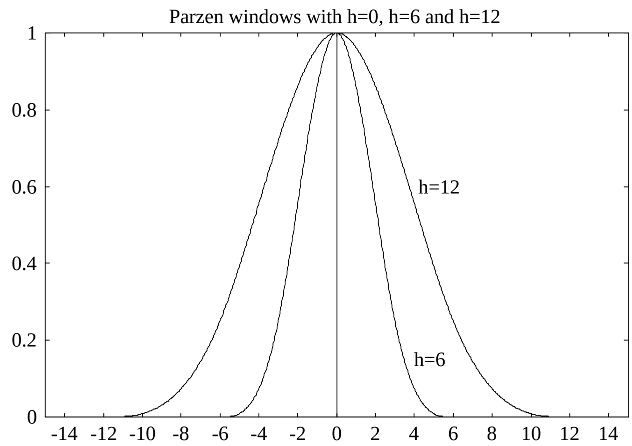
$$f\left(s_j, s_k; h\right) = \begin{cases} 1 - \frac{|j-k|}{h} & \text{if} \quad |j-k| \leq h \\ 0 & \text{otherwise} \end{cases}$$

The parameter h determines the $h-1$ unplayed strategies on either side of the played strategy whose assessments are updated. The Parzen similarity function is given by:

$$f\left(s_j, s_k; h\right) = \begin{cases} 1 - 6\left(\frac{|j-k|}{h}\right)^2 + 6\left(\frac{|j-k|}{h}\right)^3 & \text{if} \quad \left(\frac{|j-k|}{h}\right) \leq \frac{1}{2} \\ 2\left(1 - \frac{|j-k|}{h}\right)^3 & \text{if} \quad \frac{1}{2} < \left(\frac{|j-k|}{h}\right) \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

The parameter h plays a similar role here. The shapes of the Bartlett and Parzen similarity functions for these values of $h \in \{0, 6, 12\}$ are given below:





Using either similarity function would lead the player to move the assessments of all strategies in the window in the direction of the received payoff, with the assessments of strategies closer to the chosen strategy being updated more. Hence, both similarity functions seem appropriate when the strategy space is large and the agent believes that nearby strategies will give "nearby" payoffs.

In environments in which the agents know little of the decision problem, and observe little at each stage, having a similarity relation between strategies may allow the agent to utilize the information actually obtained more "efficiently". In our specification, similarity between strategies allows information regarding payoffs from played strategies to also be relevant in updating the assessments of unplayed strategies. This may represent "efficient" use of scarce information and information processing resources available to the decision maker, and represent a useful way to simplify the decision problem she faces. Of course, it may also mislead the player.

We could also have considered other similarity functions. For example, we could have supposed that the agents thought of strategies in a symmetric window around the chosen strategy to be identical. If this were assumed then f would take the value of one for nearby strategies. We could also have supposed that the similarity function was not symmetric or that nearby strategies were better than the chosen strategy (which would imply that f takes a value greater than one around the chosen strategy, assuming the payoffs to be positive). We did not consider these similarity functions because they were not appropriate for the analysis of the experimental data. For example, having $f > 1$ for some strategies would prevent choice from converging to any strategy.

3 The VBR Experiments[6]

Groups of five individuals repeatedly played the same (coordination) game for 75 periods. Each individual i had the same set of 101 pure strategies $S^i = \{0, 1, ..., 100\}$. In each period, each individual chose one strategy and received a payoff. Let s^i denote the pure strategy chosen by player i, and s^{-i} denote the strategies chosen by players other than player i. The payoff

[6] These experiments were intended to be "minimal information" treatments of Van Huyck, Cook and Battalio (1994), and were intended to be appropriate for testing reinforcement learning models (e.g. Börgers and Sarin (2000) and Erev and Roth (1998)).

function for each player i was given by

$$\pi\left(e^i, e^{-i}\right) = 0.5 - \left|e^i - \omega M(e)(1 - M(e))\right|$$

where $e^i = 1 - \frac{s^i}{100}$, $e = \left(e^1, ..., e^5\right)$, $M(e)$ is the median when the strategy profile chosen by the 5 players is e, and ω is a constant (which took the value approximately equal to 2.44 in some experiments and 3.85 in other experiments). Players were not informed about the payoff function, or the strategies available to the other players.

The game has two strict Nash equilibria which are symmetric and efficient. The equilibrium strategy of a player depends not only on the strategy the player and her opponents choose, but also on ω. We refer to the game with ω equal to approximately 2.44 as $G(2.44)$ and the game with ω approximately equal to 3.85 as $G(3.85)$. The strict equilibria of $G(2.44)$ involved each player choosing (41) and (100). For $G(3.85)$ the strict equilibria were (26) and (100). Each equilibria involves the player choosing her strategy so that $e^i = \omega M(e)(1 - M(e))$.

Every player in the game had information on (a) the number of repetitions, (b) the payoff each player would obtain would depend on the choice of that player and that of four other unidentified people in the group, (c) the payoff function was deterministic (so that if all players in the group took the same action as previously then they would receive the same payoff as previously).

The experimental results revealed that:

1. For both values of ω statistical tests could not reject the null hypothesis that initial play was uniform.

2. In all eight groups (of 5 players in each group), the median player converged to the interior equilibrium. Choices of the median player in each of the groups is illustrated in Figures 1a and 1b.

3. The four groups who played $G(2.44)$ first chose the interior equilibrium (41) between periods 10 and 18 and got "absorbed", or never left this interior equilibrium, after between periods 18 and 25. Details of the choices of each of the players in each of the groups is provided in Figure 2a.

4. The four groups who played $G(3.85)$ first chose the interior equilibrium (26) between periods 10 and 32 and got "absorbed", or never left this

interior equilibrium, after between periods 14 and 55. Details of the choices of each of the players in each of the groups is provided in Figure 2b.

5. In the final period of the stage game most players were playing the interior equilibrium strategy. Among the players who were not, they were playing strategies that earned them within 10 cents of what the equilibrium strategy paid. The configuration of the end period payoffs of the 20 subjects who played each game is given in the table below:

$G\,(2.44)$		$G\,(3.85)$	
Payoff	Frequency	Payoff	Frequency
41c	1	40c	1
46c	1	47c	3
49c	1	48c	2
50c	17	50c	14

There were two remarkable features of the experimental results. First, that in each group (there were 8 in all, with 4 playing each value of ω) play of the median player always converged to the interior equilibrium. This was remarkable given that each of the equilibria was strict, symmetric and efficient. It was hence difficult to explain why the interior equilibrium proved to be "focal". Second, the convergence to the interior equilibrium was extremely rapid. In particular, the median subject first chose the interior equilibrium and finally got absorbed in the equilibrium, well before any of the subjects could have tried more than 50% of the available strategies. This second feature led us to suspect that the labelling of the strategies, which was common across the players, was used by them to coordinate so quickly on the interior equilibrium.

4 Simulation and Estimation

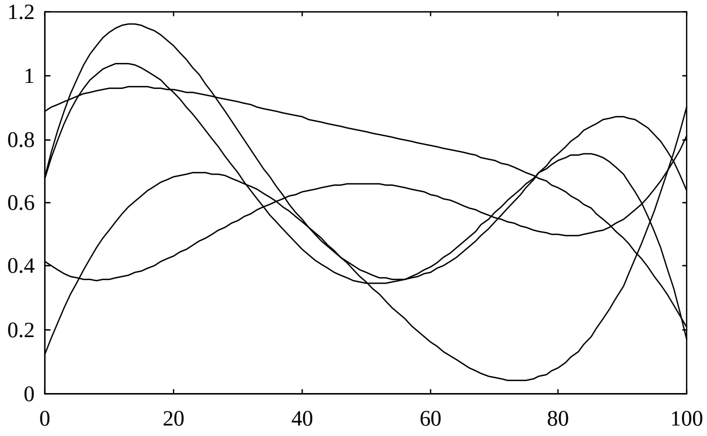
4.1 Simulations

In this subsection we investigate, by way of simulations, whether the model specified in the previous section produces median player dynamics similar to the observed median dynamics in the VBR experiment. To simulate the model we need to specify values of the initial payoff assessments of the players, values of the update parameter λ in equation (2), and the similarity functions. Given this information, and the knowledge of the payoff function used in the VBR experiment, we can simulate the model.

We considered the Bartlett and Parzen similarity functions for window size $h = 0, 6, 12$. We performed simulations for four values of the update parameter, specifically $\lambda = \{0.25, 0.50, 0.75, 1.00\}$. For the initial assessments, we consider six possible cases where five are drawn randomly from different distributions and one where the assessments lie on a random smooth polynomial. We looked at initial assessments that were (i) uniformly distributed on $[-1.00, 1.00]$; (ii) normally distributed with mean 0.25 and standard deviation 0.25; (iii) distributed according to the symmetric triangular distribution on $[-0.50, 1.00]$; (iv) uniformly distributed on $[0.50, 1.00]$; (v) uniformly distributed on $[0.00, 0.50]$; (vi) lie on a smooth polynomial.

We looked at (i), (ii) and (iii) because they seemed reasonable, given what may have been thought of as likely payoffs for participation in the experiment by the subjects, and the hint in the instructions provided to the students that payoffs might be negative. We looked at (iv) because it represented "optimistic" assessments, and such assessments played a special role in Sarin and Vahid (1999). We looked at (v) because it represented the case where the range of the assessments was roughly the same as the range of the actual payoffs in the experiment. Lastly, in (vi) we considered initial assessments that embodied the idea that similar strategies also have similar initial assessments. This was implemented by choosing assessments of 6 strategies 0, 20, 40, 60, 80, 100 independently from a uniform distribution on $[0, 1]$, and then finding the polynomial of degree 5 which passed through these points. The assessments of the 101 strategies were then read off this polynomial. An example of such a configuration of initial assessments is given in the following figure:

An example of initial assessments of five players according to (vi)



To summarize, the design space of the simulations is the cross product of six distributions of initial assessments, five similarity functions (both Bartlett and Parzen functions with zero window width degenerate to the same case in which there is no similarity among the actions), and four update parameters for each of the two games $G(2.44)$ and $G(3.85)$. To ensure that any particular draw of the random initial assessments is not given too much weight in our conclusions, we perform 100 simulations for each case.

Tables 1.a through 1.d present the 5th percentile, the mode and the 95th percentile of the empirical distribution of the 100 simulated medians at the 25th, 50th and the final period of the game. We present these particular summary statistics because in the actual experiments, the median choice settled at the interior equilibrium after the 25th period for all four cohorts playing $G(2.44)$, and after the 50th period for all four cohorts playing $G(3.85)$ (see Figures 1.a and 1.b).

Looking at the $h = 0$ results, we see that for either game, the performance of the payoff assessment model without similarity is much better than the Cross model in explaining the data. However, the modal value in many of the reported cases is not at the interior equilibrium, and when it is, the distribution of the median is quite dispersed around this modal value. The only case where the simulated medians were all in a close proximity of the interior equilibrium is the case where the support of the initial assessments is close to their actual support, i.e. case (v) above. Different values of λ do not change the conclusions significantly.

For other values of h, the simulations show that increasing λ increases the height of the empirical distribution at the modal value. Increasing h will pull the modal value to the interior equilibrium at first, but as h becomes too large, the dispersion of final median choices around the modal value increases. This can be seen from the representative histograms of the simulated medians at the final period of the games that are presented in Figure 3. Almost all distributions of initial assessments, with $\lambda \geq 0.75$ and $h \geq 6$, seem to produce median dynamics that resemble that the observed data quite closely. The exceptions are the symmetric triangular initial assessments (case (iii)) and the "optimistic" initial assessments (case (iv)), which need λ close to 1 and h close to 12 before they produce median dynamics that look like the observed ones.

Figure 4 provides a typical simulated path for the case where $h = 6$ and $\lambda = 0.75$ and uniform initial assessments distributed on $[-1, 1]$. What is remarkable is not only that the median dynamics converge to the same equilibrium selected in the experiments, but, more importantly, that this convergence takes place in roughly the same number of repetitions as takes in the experiment.[7]

4.2 Estimation

In this subsection, we provide a formal statistical test of whether the payoff distributions implied by our estimated model can be distinguished from the payoff distributions observed in the experiment. We apply the test at each period of the game.

The simulations of the previous subsection revealed that the implications of the model are robust with respect to reasonable variations in the initial assessments. This is fortunate, since the data reveals very little information about the initial assessments, especially for the assessments of the strategies that are never played. Hence, we assume that the initial assessments for each strategy and each player were drawn independently from a uniform distribution over $[-1, 1]$, and we fix the similarity function to be the Parzen similarity function.

We concentrate on the estimation of the values of λ and h which best fit the data over the entire course of the experiment under our states assump-

[7] Although VBR obtained convergence to the interior equilibrium in their simulations of the Cross model of reinforcement learning for a sufficiently slow learning rate parameter, this convergence took a very large number of repetitions. This slow convergence is noted also in other work on reinforcement learning models (e.g. Chen and Tang (1998), Possajennikov (1997)). It is also known to afflict models of evolutionary game theory (e.g. Kandori, Mailath and Rob (1993)).

tions, and choose the (λ, h) pair which explains the data best with respect to the criterion that we will explain below. Even though, for a particular value of h, there is only one unknown parameter, the maximum likelihood estimation of this parameter is not feasible. Since at every stage, there are 101 possible choices, and the game is repeated 75 times, the likelihood of an observed sequence of choices implied by our structural model involves evaluation of many-fold integrals. This makes the evaluation of the likelihood, even by simulation method, for a single value of λ computationally infeasible. Hence, we estimate λ by matching the moments of the simulated and real data.[8]

For a given h, the parameter λ is chosen to minimize the distance between the expected path of the payoffs from the theoretical model to the expected path of the actual payoffs. The expected path of the model is approximated by averaging 500 simulated paths, and the average of the observed payoff paths are taken as an estimate of the expected path of the actual payoffs. The estimate of λ thus minimizes:

$$Q(h,\lambda) = \sum_{t=1}^{75} \left(y_t^{(\omega_1)} - \tilde{y}_t^{(\omega_1)}(\lambda) \right)^2 + \sum_{t=1}^{75} \left(y_t^{(\omega_2)} - \tilde{y}_t^{(\omega_2)}(\lambda) \right)^2$$

where the superscripts (ω_1) and (ω_2) indicate the two different games that the subjects were playing, $y_t = \frac{1}{20}\sum_{j=1}^{20} \pi_{jt}$ is the average of the observed payoffs at time t, and $\tilde{y}_t(\lambda) = \frac{1}{500}\sum_{j=1}^{500} \widetilde{\pi}_{jt}(\lambda)$ is the average of the simulated payoffs at time t of the respective game. Since, given h, there is only one parameter λ to be estimated, and since λ is restricted to be between zero and one, estimating λ by a grid search is appropriate.

Figure 5 shows the value of the objective function at $\{0.60, 0.61, ..., 1.00\} \times \{1, 2, ..., 20\}$. As it can be seen from this graph, the objective function improves as h increases for almost all λ and the global minimum of the objective function is achieved when $(\lambda, h) = (.71, 18)$. It is evident that the objective function flattens out as h becomes large. The almost flatness of the objective function for $h \geq 10$ and $0.5 \leq \lambda \leq 1.0$ somewhat assures us that the assumption of identical updating parameters and similarity weights across the population is not too costly, in the sense that the simplified model is almost observationally equivalent to a model where the subjects have different h and λ, as long as for all subjects $h \geq 10$ and $0.5 \leq \lambda \leq 1.0$.

[8]This estimation method is known as the method of simulated moments. See, e.g., Gregory and Smith (1993) for a survey.

Under the assumption of correct specification, the estimator of λ which minimizes the above objective function will be consistent as long as the number of simulations is always a multiple of the sample size (i.e. it grows to infinity at least at the same rate). However, strong dependence and non-stationarity of the distribution of observations over time, makes the development of the classical tests of model adequacy very difficult. Hence, we follow VBR and use Kolmogorov-Smirnov tests (E.g. Berry and Lindgren (1996) p.534-535) of whether the observed and simulated data could have come from the same (unconditional) distributions. While the rejection of such test at the terminal stage of the game, a stage in which most subjects have converged to repeating the same strategy, is sufficiently strong for the dismissal of the theoretical model, not rejecting the test at just one stage is only weak evidence for it. We therefore perform this test for *every* stage of the game.[9]

Table 2 reports the p-values of the Kolmogorov-Smirnov test of the null hypothesis that the observed and the simulated strategy choices come from the same distribution for every stage of the game. Overall the null of equality of CDF's of the observed and simulated individual choices is rejected for 4 out of 75 periods in $G(3.85)$ and 2 out of 75 for the $G(2.44)$ at 5% level of significance.[10] We take this as strong evidence that our simple model of choice, and the postulated rule for updating the assessment of payoffs of similar strategies, can describe the behavior of decision makers who have very limited knowledge.[11]

[9] VBR performed the Kolmogorov-Smirnov test only for period 75. They rejected the hypothesis of equality between the observed distrbution and those obtained from simulations of the Cross model. In our case, the test does not reject for the hypothesis of equality of distributions when we consider only period 75. This is why we perform the stronger test.

[10] The test statistics for different periods are not independent, and, therefore, the overall level of significance will be different from 5%.

[11] We also calculated the best value of λ assuming that there is no similarity ($h = 0$) between the strategies. We think this is a useful exercise as our simulations reveal that the model without similarity can also mimic the data rather well for some initial assessments. We find that the best value of $\lambda = 0.737$ and the Kolmogorov-Smirnov tests rejected equality of distributions for 52 periods in $G(3.85)$ and for 12 periods in $G(2.44)$. These results confirm that adding similarity in the model considerably improves its explanatory power.

5 The Cross Model and Similarity

In the previous section we saw that allowing for similarity gave us very favorable results. A natural question to ask is whether incorporating considerations of similarity among the strategies in the Cross model would improve its convergence properties. In this section we modify the Cross model to allow the agent to take into account the similarity among strategies. Another modification of the Cross model we consider allows for a declining step-size or learning rate over time.

The Cross model assumes that an agent has a finite set $S = \{s_1, ..., s_J\}$ of strategies. The agent is assumed to know S but not the payoff function she faces. At time n she chooses among her strategies according a $J \times 1$ probability vector $p(n)$. Upon choosing a strategy she receives a payoff. She uses this payoff to update the vector p. If at time n she chooses strategy s_j and receives a payoff of π_j then she updates her probability vector as follows:

$$p(n+1) = p(n) + \alpha \pi_j (e_j - p(n)) \tag{3}$$

where $\alpha \in (0, 1]$ is the "step size" or learning rate parameter and e_j is a $J \times 1$ unit vector with a one in the $j-th$ position and zeroes elsewhere. For the Cross learning rule to be well-defined we require "effective payoffs" $\alpha \pi_j$ to be normalized to lie between zero and one.

VBR performed an extensive set of simulations with the Cross model. They assumed the initial mixed strategy vector $p(0)$ to be uniformly distributed over the 101 strategies available to each player. Simulations were performed for the two payoff functions ($G(2.44)$, and $G(3.85)$) and for different values of α. For $\alpha \geq 0.1$ they observed that the Cross model did not necessarily converge to the interior equilibrium. For $\alpha \leq 0.05$ they observed that the Cross model did usually converge to the interior equilibrium. This convergence, however, took a very large number of repetitions. For $\alpha = 0.05$, they observed that the Cross model converged in about 750 repetitions. This was very different from the experimental results in which subjects converged in less than 50 repetitions.

To answer whether introducing similarity relations into the Cross model would help in explaining the data we performed a series of simulations with a modified version of the Cross model in which the agent shifts probability mass to all similar strategies. As in the previous section, we suppose that similar strategies are those that lie within a certain window around the played strategy. This is achieved by replacing the unit vector e_j with a vector whose components sum to one, and whose components decrease in value as they get further from the played strategy, and which has zeroes elsewhere. Specifically, e_j in (3) is replaced by the $J \times 1$ vector d_j which has in its $k-th$ row the term $f(s_j, s_k) / \sum_l f(s_j, s_l)$. The numerator is the similarity function defined in Section 3, and the denominator is required because probabilities need to sum to one. Formally, the Cross model with similarity is given by,

$$p(n+1) = p(n) + \alpha \pi_j (d_j - p(n)) \tag{4}$$

where d_j is defined above.

As before, we considered different window lengths and allowed for different values of α. Our simulations revealed that the convergence property of the Cross model was destroyed altogether. That is, the Cross model with similarity did not converge to the interior equilibrium. Figure 6.a gives the histogram of the median choices in periods 75 and 750 in a hundred simulations of the modified Cross model using the Parzen similarity function for $G(2.44)$, and figure 6.b gives the analogous diagrams for $G(3.85)$. As is readily seen, supposing that players update the assessments of nearby strategies in a similar way worsens the convergence properties of the Cross dynamic in both games. The intuition for this result is that because the agent updates similar but unplayed strategies the probability of these strategies does not decline as is assumed in the Cross model. This prevents convergence of the median choice to the interior equilibrium, even when learning is slow.

We next consider another modification of the Cross model. We consider the case in which the learning rate parameter, or step-size parameter, of the Cross model declines over time. Specifically, we consider the case where the step size decreases at rate $1/n$. We hoped that such a modification of the Cross model might improve its convergence properties, with or without incorporating similarity among the strategies. Our simulations, reported in Figures 6.c and 6.d, reveal that this is not the case. An inspection of these figures shows that the convergence properties of the model deteriorate with a declining step-size, and that this is true with or without the consideration of similarity among the alternate strategies.

Our conclusion in this section is that modifying the Cross model to allow for either similarity among strategies or a decreasing learning rate parameter does not improve the predictive model of the Cross model. Whereas the Cross learning rule converges to the equilibrium selected by the subjects in the VBR experiments, even while it takes a much longer time to do so, modifying the Cross dynamic in the above mentioned ways makes it not converge to any equilibrium at all.

6 Conclusion

In this paper we have incorporated similarity, that people may perceive among strategies, in a simple dynamic model of choice. Decision makers may assume similarity among strategies to be present in many economic situations, especially when the set of strategies is large. Using similarity relations may help make the strategy space "manageable". We introduced two similarity functions, the Parzen similarity function and the Bartlett similarity function. Each has the property that strategies which are more distant are viewed by the decision maker as less similar. Many strategy spaces have this property (e.g. strategy spaces of output and price). Another important characteristic of the similarity relations we introduce is that they involve the addition of only one additional parameter and hence allow for a parsimonious model specification.

We have shown that incorporating similarity improves the explanatory power of the model. The model with similarity that we developed not only converged to the same equilibrium chosen by the experimental subjects, in a situation where traditional refinements could not distinguish among the equilibria, but also converged to the equilibrium in roughly the same number of repetitions. This, and the fact that formal statistical tests could not distinguish between the payoff distributions of the simulated model and the observed payoff distributions in almost every period provided striking support for the model.

In Sarin and Vahid (2001) we have shown that the payoff assessment model, without similarity, can explain data better than reinforcement learning models in zero-sum games. In this paper we have shown that the same model, with or without similarity, outperforms reinforcement learning models in coordination games. Hence, we have obtained further evidence in favor of the payoff assessment model.

References

[1] Berry, D. and B. Lindgren (1996): *Statistics: Theory and Methods*, 2nd edition, Duxbury Press.

[2] Börgers, T., A. Morales and R. Sarin (2001): Expedient and monotone learning rules," mimeo, UCL, Malaga, TAMU.

[3] Börgers, T. and R. Sarin (2000): "Naive Reinforcement Learning with Endogenous Aspirations," *International Economic Review*, 41, 921-950.

[4] Brockwell, P. and R. Davis (1991): *Time Series: Theory and Methods*, 2nd edition, Springer-Verlag.

[5] Chen, Y. and F. Tang (1998), "Learning and Incentive Compatible Mechanisms for Public Goods Provision: An Experimental Study," *Journal of Political Economy*, 106, 633-662.

[6] Cross, J. (1973): "A Stochastic Learning Model of Economic Behavior," *Quarterly Journal of Economics*, 87, 239-266.

[7] Erev, I. and A. Roth (1998): "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibrium," *American Economic Review*, 88, 848-881.

[8] Gilboa, I. and D. Schmeidler (1995): "Case-Based Decision Theory," *Quarterly Journal of Economics*, 110, 605-639.

[9] Gilboa, I. and D. Schmeidler (1997): "Act Similarity in Case-Based Decision Theory," *Economic Theory*, 9, 47-61.

[10] Gregory, A.W. and G.W. Smith (1993): "Statistical Aspects of Calibration in Macroeconomics," Chapter 25 in G.S. Maddala, C.R. Rao and H.D. Vinod (eds), *Handbook of Statistics, Volume II: Econometrics*, North Holland.

[11] Huck, S., H. Norman and J. Oechssler (1999): "Learning in Cournot oligopoly: An experiment," *Economic Journal*, 109, C80-C95.

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

[13] Possajennikov, A. (1997): "An Analysis of a Simple Reinforcing Dynamic: Learning to Play an "Egalitarian" Equilibrium," Mimeo, Tilburg University.

[14] Rubinstein, A. (1988): "Similarity and Decision-Making under Risk (Is There a Utility Theory Resolution to the Allais Paradox)," *Journal of Economic Theory*, 46, 145-153.

[15] Sarin, R. and F. Vahid (1999): "Payoff Assessment without Probabilities: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, 28, 294-309.

[16] Sarin, R. and F. Vahid (2001): "Predicting How People Play Games: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, 34, 104-122.

[17] Van Huyck, J., J. Cook, and R. Battalio (1994): "Selection Dynamics, Asymptotic Stability, and Adaptive Behavior," *Journal of Political Economy*, 102, 975-1005.

[18] Van Huyck, J., R. Battalio and F. Rankin (1996): "Selection Dynamics and Adaptive Behavior without much Information," Mimeo, Texas A&M University.

Table 1.a: (5th percentile,mode,95th percentile) of the distribution of median in 100 simulations of G(2.44) with Bartlett similarity function
(*0.59 is the interior equilibrium of G(2.44)*)

Distribution of initial assessments	$\lambda = 0.25$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.22,.46,.77)	(.29,.56,.65)	(.37,.59,.71)	(.27,.57,.69)	(.38,.59,.64)	(.57,.59,.61)	(.41,.59,.71)	(.56,.59,.61)	(.57,.59,.60)
N(0.25,0.25^2)	(.28,.58,.72)	(.54,.59,.61)	(.57,.59,.61)	(.44,.58,.69)	(.57,.59,.61)	(.58,.59,.60)	(.49,.59,.63)	(.57,.59,.61)	(.57,.59,.60)
STRI(-0.5,1)	(.15,.63,.80)	(.27,.62,.81)	(.17,.61,.80)	(.24,.48,.80)	(.24,.59,.74)	(.24,.60,.74)	(.16,.44,.81)	(.27,.59,.73)	(.58,.59,.61)
SM(0,1,5)	(.13,.58,.86)	(.14,.31,.76)	(.26,.49,.76)	(.30,.59,.76)	(.43,.59,.65)	(.52,.59,.60)	(.48,.59,.66)	(.56,.59,.61)	(.56,.59,.61)
U(0.5,1)	(.17,.39,.80)	(.22,.65,.76)	(.20,.51,.75)	(.23,.48,.73)	(.29,.58,.77)	(.43,.59,.66)	(.24,.53,.75)	(.57,.59,.61)	(.57,.59,.60)
U(0,0.5)	(.54,.58,.62)	(.56,.58,.62)	(.56,.58,.62)	(.55,.58,.62)	(.57,.58,.61)	(.57,.58,.61)	(.55,.58,.62)	(.57,.58,.61)	(.57,.58,.61)

Distribution of initial assessments	$\lambda = 0.50$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.26,.49,.71)	(.54,.59,.63)	(.57,.59,.60)	(.49,.60,.66)	(.57,.59,.60)	(.57,.59,.60)	(.57,.59,.61)	(.57,.59,.60)	(.52,.59,.61)
N(0.25,0.25^2)	(.39,.60,.66)	(.58,.59,.61)	(.57,.59,.61)	(.57,.59,.61)	(.58,.59,.60)	(.58,.59,.60)	(.56,.59,.61)	(.57,.59,.61)	(.54,.59,.61)
STRI(-0.5,1)	(.15,.33,.80)	(.25,.54,.66)	(.16,.27,.81)	(.20,.50,.77)	(.29,.58,.69)	(.58,.59,.60)	(.28,.61,.76)	(.57,.59,.60)	(.58,.59,.60)
SM(0,1,5)	(.20,.60,.75)	(.21,.53,.72)	(.29,.53,.69)	(.48,.59,.63)	(.56,.59,.60)	(.56,.59,.60)	(.52,.59,.61)	(.54,.59,.61)	(.50,.59,.60)
U(0.5,1)	(.21,.25,.81)	(.21,.45,.75)	(.31,.52,.74)	(.26,.63,.75)	(.58,.59,.60)	(.58,.59,.59)	(.53,.59,.62)	(.57,.59,.60)	(.57,.59,.60)
U(0,0.5)	(.55,.58,.62)	(.55,.58,.62)	(.55,.58,.62)	(.56,.58,.62)	(.56,.58,.62)	(.56,.58,.62)	(.56,.58,.61)	(.57,.58,.61)	(.57,.59,.61)

Distribution of initial assessments	$\lambda = 0.75$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.32,.60,.65)	(.57,.59,.60)	(.57,.59,.60)	(.57,.59,.61)	(.57,.59,.60)	(.58,.59,.60)	(.56,.59,.60)	(.56,.59,.61)	(.56,.59,.60)
N(0.25,0.25^2)	(.51,.59,.62)	(.57,.59,.61)	(.57,.59,.61)	(.57,.59,.60)	(.56,.59,.62)	(.58,.59,.60)	(.56,.59,.61)	(.57,.59,.61)	(.57,.59,.63)
STRI(-0.5,1)	(.15,.33,.80)	(.32,.59,.63)	(.26,.61,.71)	(.23,.58,.76)	(.58,.59,.60)	(.57,.59,.62)	(.57,.59,.60)	(.57,.59,.60)	(.56,.59,.61)
SM(0,1,5)	(.26,.60,.75)	(.26,.59,.73)	(.40,.59,.63)	(.53,.59,.61)	(.57,.59,.60)	(.58,.59,.60)	(.54,.59,.61)	(.57,.59,.61)	(.57,.59,.62)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.15,.61,.81)	(.22,.57,.64)	(.58,.59,.59)	(.58,.59,.60)	(.56,.59,.61)	(.56,.59,.61)	(.57,.59,.60)
U(0,0.5)	(.54,.58,.62)	(.54,.58,.62)	(.54,.58,.62)	(.56,.59,.60)	(.56,.59,.61)	(.58,.59,.60)	(.55,.58,.60)	(.56,.59,.60)	(.56,.59,.61)

Distribution of initial assessments	$\lambda = 1.00$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.25,.59,.72)	(.56,.59,.60)	(.56,.59,.60)	(.56,.59,.60)	(.58,.59,.60)	(.58,.59,.60)	(.56,.59,.61)	(.56,.59,.61)	(.53,.58,.61)
N(0.25,0.25^2)	(.50,.58,.62)	(.56,.58,.61)	(.56,.58,.61)	(.57,.59,.60)	(.58,.59,.60)	(.59,.59,.60)	(.52,.59,.66)	(.56,.59,.60)	(.51,.59,.60)
STRI(-0.5,1)	(.15,.33,.80)	(.34,.58,.61)	(.58,.59,.60)	(.25,.59,.69)	(.55,.59,.62)	(.54,.59,.60)	(.56,.59,.64)	(.57,.59,.65)	(.56,.58,.61)
SM(0,1,5)	(.26,.60,.75)	(.28,.59,.73)	(.50,.59,.60)	(.55,.59,.60)	(.58,.59,.60)	(.58,.59,.60)	(.54,.59,.63)	(.56,.59,.61)	(.57,.59,.61)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.15,.66,.82)	(.56,.59,.61)	(.58,.59,.60)	(.54,.59,.60)	(.57,.59,.61)	(.50,.59,.62)	(.54,.59,.60)
U(0,0.5)	(.54,.58,.62)	(.54,.58,.62)	(.54,.58,.62)	(.55,.59,.61)	(.58,.59,.61)	(.58,.59,.62)	(.53,.59,.63)	(.53,.59,.64)	(.53,.59,.61)

Table 1.b: (5th percentile,mode,95th percentile) of the distribution of median in 100 simulations of G(3.85) with Bartlett similarity function
(0.74 is the interior equilibrium of G(3.85))

Distribution of initial assessments	$\lambda = 0.25$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.26,.76,.85)	(.28,.74,.84)	(.49,.75,.84)	(.27,.74,.85)	(.52,.74,.81)	(.73,.74,.75)	(.37,.73,.87)	(.72,.74,.76)	(.72,.74,.76)
N(0.25,0.25^2)	(.35,.73,.85)	(.54,.74,.78)	(.72,.74,.76)	(.43,.78,.85)	(.72,.74,.76)	(.72,.74,.75)	(.63,.74,.79)	(.72,.74,.76)	(.72,.74,.76)
STRI(-0.5,1)	(.15,.33,.80)	(.23,.70,.83)	(.21,.58,.83)	(.24,.48,.85)	(.32,.73,.87)	(.38,.67,.88)	(.23,.51,.88)	(.40,.65,.84)	(.72,.74,.76)
SM(0,1,5)	(.16,.33,.88)	(.22,.64,.85)	(.37,.66,.81)	(.36,.76,.88)	(.72,.74,.77)	(.73,.74,.76)	(.66,.74,.80)	(.72,.74,.76)	(.72,.74,.76)
U(0.5,1)	(.18,.25,.81)	(.30,.61,.83)	(.30,.53,.80)	(.31,.48,.86)	(.31,.70,.83)	(.58,.74,.76)	(.38,.75,.88)	(.71,.74,.76)	(.72,.74,.75)
U(0,0.5)	(.53,.75,.79)	(.72,.75,.76)	(.72,.75,.76)	(.72,.73,.76)	(.72,.73,.76)	(.72,.73,.76)	(.71,.75,.76)	(.72,.74,.76)	(.72,.74,.76)

Distribution of initial assessments	$\lambda = 0.50$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.29,.71,.88)	(.56,.74,.80)	(.72,.74,.76)	(.48,.75,.79)	(.72,.74,.76)	(.72,.74,.76)	(.71,.74,.77)	(.72,.74,.76)	(.72,.74,.77)
N(0.25,0.25^2)	(.27,.73,.84)	(.71,.74,.76)	(.72,.74,.76)	(.71,.73,.76)	(.73,.73,.76)	(.73,.73,.76)	(.71,.74,.76)	(.71,.74,.76)	(.65,.74,.78)
STRI(-0.5,1)	(.15,.33,.80)	(.34,.66,.90)	(.30,.54,.79)	(.22,.75,.85)	(.49,.73,.83)	(.73,.74,.75)	(.33,.75,.85)	(.72,.74,.76)	(.72,.74,.76)
SM(0,1,5)	(.16,.60,.79)	(.21,.67,.85)	(.43,.72,.82)	(.65,.74,.78)	(.73,.74,.76)	(.73,.74,.76)	(.71,.75,.76)	(.71,.74,.76)	(.67,.74,.77)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.79)	(.20,.71,.84)	(.41,.53,.82)	(.72,.74,.76)	(.72,.74,.75)	(.66,.73,.77)	(.71,.74,.76)	(.71,.74,.76)
U(0,0.5)	(.49,.75,.80)	(.72,.73,.77)	(.72,.73,.77)	(.71,.74,.76)	(.72,.74,.76)	(.72,.74,.76)	(.72,.74,.76)	(.72,.74,.76)	(.71,.74,.76)

Distribution of initial assessments	$\lambda = 0.75$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.37,.56,.90)	(.71,.74,.76)	(.72,.74,.76)	(.72,.74,.77)	(.72,.74,.76)	(.73,.74,.75)	(.71,.74,.76)	(.66,.74,.79)	(.71,.74,.75)
N(0.25,0.25^2)	(.33,.72,.83)	(.71,.74,.76)	(.71,.74,.76)	(.72,.74,.76)	(.72,.74,.76)	(.73,.74,.75)	(.71,.74,.76)	(.70,.74,.81)	(.71,.74,.76)
STRI(-0.5,1)	(.15,.33,.80)	(.32,.72,.87)	(.54,.73,.80)	(.32,.61,.86)	(.72,.74,.76)	(.69,.74,.77)	(.68,.74,.78)	(.71,.75,.76)	(.72,.74,.76)
SM(0,1,5)	(.26,.60,.77)	(.29,.71,.87)	(.41,.72,.83)	(.72,.74,.78)	(.72,.74,.76)	(.73,.74,.75)	(.71,.74,.76)	(.68,.74,.78)	(.72,.73,.83)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.17,.43,.81)	(.37,.73,.83)	(.72,.74,.75)	(.73,.74,.75)	(.70,.74,.76)	(.67,.74,.81)	(.72,.74,.76)
U(0,0.5)	(.33,.74,.84)	(.72,.75,.77)	(.72,.75,.77)	(.71,.74,.76)	(.71,.74,.77)	(.73,.74,.75)	(.71,.74,.76)	(.71,.74,.76)	(.67,.74,.81)

Distribution of initial assessments	$\lambda = 1.00$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.25,.73,.89)	(.71,.74,.78)	(.71,.74,.78)	(.71,.74,.77)	(.73,.74,.76)	(.73,.74,.75)	(.70,.74,.76)	(.64,.74,.80)	(.69,.73,.75)
N(0.25,0.25^2)	(.22,.75,.88)	(.71,.73,.77)	(.71,.75,.77)	(.72,.74,.77)	(.73,.74,.76)	(.73,.74,.75)	(.70,.74,.77)	(.65,.73,.82)	(.72,.73,.77)
STRI(-0.5,1)	(.15,.33,.80)	(.32,.73,.86)	(.72,.73,.76)	(.47,.74,.85)	(.71,.75,.76)	(.69,.74,.78)	(.69,.74,.77)	(.70,.74,.76)	(.71,.73,.80)
SM(0,1,5)	(.24,.60,.78)	(.35,.71,.85)	(.54,.74,.83)	(.72,.74,.77)	(.73,.74,.76)	(.73,.74,.75)	(.68,.74,.77)	(.67,.74,.76)	(.71,.73,.76)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.15,.66,.82)	(.68,.74,.77)	(.71,.74,.76)	(.73,.74,.75)	(.70,.74,.77)	(.64,.73,.80)	(.72,.73,.77)
U(0,0.5)	(.31,.74,.82)	(.71,.74,.77)	(.71,.74,.77)	(.71,.74,.76)	(.73,.74,.75)	(.73,.74,.75)	(.66,.75,.82)	(.66,.73,.81)	(.72,.75,.75)

Table 1.c: (5th percentile,mode,95th percentile) of the distribution of median in 100 simulations of G(2.44) with Parzen similarity function

(*0.59 is the interior equilibrium of G(2.44)*)

Distribution of initial assessments	$\lambda = 0.25$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.22,.46,.77)	(.29,.56,.65)	(.37,.59,.71)	(.28,.56,.76)	(.36,.58,.71)	(.56,.59,.61)	(.29,.56,.77)	(.54,.59,.61)	(.57,.59,.60)
N(0.25,0.25^2)	(.28,.58,.72)	(.54,.59,.61)	(.57,.59,.61)	(.33,.59,.69)	(.57,.59,.61)	(.58,.59,.60)	(.47,.59,.64)	(.57,.59,.61)	(.58,.59,.60)
STRI(-0.5,1)	(.15,.63,.80)	(.27,.62,.81)	(.17,.61,.80)	(.24,.53,.78)	(.21,.53,.80)	(.25,.39,.79)	(.19,.43,.77)	(.23,.58,.79)	(.31,.58,.67)
SM(0,1,5)	(.13,.58,.86)	(.14,.31,.76)	(.26,.49,.76)	(.30,.52,.71)	(.33,.59,.66)	(.53,.59,.61)	(.35,.59,.65)	(.53,.59,.61)	(.54,.59,.60)
U(0.5,1)	(.17,.39,.80)	(.22,.65,.76)	(.20,.51,.75)	(.24,.46,.77)	(.32,.63,.69)	(.33,.57,.72)	(.19,.39,.73)	(.41,.58,.66)	(.58,.59,.60)
U(0,0.5)	(.54,.58,.62)	(.56,.58,.62)	(.56,.58,.62)	(.55,.58,.62)	(.56,.58,.62)	(.56,.58,.62)	(.55,.58,.62)	(.55,.58,.62)	(.55,.58,.62)

Distribution of initial assessments	$\lambda = 0.50$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.26,.49,.71)	(.54,.59,.63)	(.57,.59,.60)	(.34,.57,.71)	(.57,.59,.60)	(.58,.59,.60)	(.54,.59,.63)	(.57,.59,.60)	(.57,.59,.60)
N(0.25,0.25^2)	(.39,.60,.66)	(.58,.59,.61)	(.57,.59,.61)	(.55,.59,.61)	(.58,.59,.60)	(.58,.59,.60)	(.57,.59,.61)	(.57,.59,.60)	(.58,.59,.60)
STRI(-0.5,1)	(.15,.33,.80)	(.25,.54,.66)	(.16,.27,.81)	(.18,.54,.74)	(.34,.51,.71)	(.53,.59,.61)	(.22,.54,.77)	(.57,.59,.61)	(.58,.59,.60)
SM(0,1,5)	(.20,.60,.75)	(.21,.53,.72)	(.29,.53,.69)	(.36,.59,.71)	(.55,.59,.60)	(.55,.59,.60)	(.51,.59,.61)	(.55,.59,.60)	(.57,.59,.60)
U(0.5,1)	(.21,.25,.81)	(.21,.45,.75)	(.31,.52,.74)	(.25,.51,.71)	(.40,.59,.65)	(.58,.59,.59)	(.28,.57,.69)	(.58,.59,.60)	(.58,.59,.59)
U(0,0.5)	(.55,.58,.62)	(.55,.58,.62)	(.55,.58,.62)	(.56,.58,.61)	(.56,.59,.61)	(.56,.59,.61)	(.56,.58,.61)	(.56,.59,.61)	(.58,.59,.60)

Distribution of initial assessments	$\lambda = 0.75$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.32,.60,.65)	(.57,.59,.60)	(.57,.59,.60)	(.56,.59,.61)	(.57,.59,.60)	(.59,.59,.59)	(.56,.59,.61)	(.58,.59,.61)	(.57,.59,.60)
N(0.25,0.25^2)	(.51,.59,.62)	(.57,.59,.61)	(.57,.59,.61)	(.56,.59,.60)	(.57,.59,.60)	(.58,.59,.60)	(.57,.59,.60)	(.58,.59,.60)	(.58,.59,.60)
STRI(-0.5,1)	(.15,.33,.80)	(.32,.59,.63)	(.26,.61,.71)	(.21,.41,.77)	(.57,.59,.60)	(.58,.59,.59)	(.33,.58,.65)	(.58,.59,.60)	(.58,.59,.60)
SM(0,1,5)	(.26,.60,.75)	(.26,.59,.73)	(.40,.59,.63)	(.50,.59,.64)	(.56,.59,.60)	(.58,.59,.60)	(.55,.59,.61)	(.55,.59,.61)	(.55,.59,.60)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.15,.61,.81)	(.32,.58,.74)	(.58,.59,.59)	(.56,.59,.60)	(.56,.59,.60)	(.54,.59,.61)	(.58,.59,.60)
U(0,0.5)	(.54,.58,.62)	(.54,.58,.62)	(.54,.58,.62)	(.54,.59,.61)	(.57,.59,.61)	(.57,.59,.60)	(.56,.59,.61)	(.58,.59,.60)	(.58,.59,.60)

Distribution of initial assessments	$\lambda = 1.00$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.25,.59,.72)	(.56,.59,.60)	(.56,.59,.60)	(.56,.59,.60)	(.58,.59,.60)	(.59,.59,.59)	(.55,.59,.62)	(.58,.59,.60)	(.53,.59,.59)
N(0.25,0.25^2)	(.50,.58,.62)	(.56,.58,.61)	(.56,.58,.61)	(.57,.59,.60)	(.59,.59,.60)	(.56,.59,.61)	(.54,.59,.63)	(.56,.59,.64)	(.53,.59,.59)
STRI(-0.5,1)	(.15,.33,.80)	(.34,.58,.61)	(.58,.59,.60)	(.21,.38,.75)	(.57,.59,.60)	(.56,.59,.62)	(.57,.59,.61)	(.53,.59,.59)	(.53,.59,.64)
SM(0,1,5)	(.26,.60,.75)	(.28,.59,.73)	(.50,.59,.60)	(.53,.59,.61)	(.56,.59,.61)	(.58,.59,.59)	(.55,.59,.62)	(.53,.59,.63)	(.53,.59,.59)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.15,.66,.82)	(.29,.56,.69)	(.58,.59,.60)	(.56,.59,.59)	(.57,.59,.60)	(.55,.59,.60)	(.53,.59,.59)
U(0,0.5)	(.54,.58,.62)	(.54,.58,.62)	(.54,.58,.62)	(.55,.59,.61)	(.58,.59,.61)	(.56,.59,.61)	(.57,.59,.61)	(.53,.59,.64)	(.58,.59,.59)

Table 1.d: (5th percentile,mode,95th percentile) of the distribution of median in 100 simulations of G(3.85) with Parzen similarity function
(0.74 is the interior equilibrium of G(3.85))

Distribution of initial assessments	$\lambda = 0.25$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.26,.76,.85)	(.28,.74,.84)	(.49,.75,.84)	(.31,.53,.88)	(.52,.73,.83)	(.72,.74,.76)	(.36,.68,.87)	(.70,.74,.76)	(.72,.74,.75)
N(0.25,0.25^2)	(.35,.73,.85)	(.54,.74,.78)	(.72,.74,.76)	(.40,.68,.84)	(.72,.74,.76)	(.72,.74,.75)	(.44,.73,.81)	(.72,.74,.76)	(.72,.74,.76)
STRI(-0.5,1)	(.15,.33,.80)	(.23,.70,.83)	(.21,.58,.83)	(.21,.64,.80)	(.18,.67,.85)	(.25,.48,.84)	(.24,.58,.83)	(.28,.69,.86)	(.49,.74,.81)
SM(0,1,5)	(.16,.33,.88)	(.22,.64,.85)	(.37,.66,.81)	(.31,.73,.83)	(.60,.74,.78)	(.72,.74,.75)	(.55,.74,.80)	(.72,.74,.76)	(.73,.74,.76)
U(0.5,1)	(.18,.25,.81)	(.30,.61,.83)	(.30,.53,.80)	(.37,.71,.89)	(.31,.72,.83)	(.39,.75,.84)	(.34,.75,.89)	(.49,.74,.80)	(.73,.74,.76)
U(0,0.5)	(.53,.75,.79)	(.72,.75,.76)	(.72,.75,.76)	(.70,.75,.76)	(.72,.75,.76)	(.72,.75,.76)	(.72,.74,.76)	(.72,.74,.76)	(.72,.74,.76)
Distribution of initial assessments	$\lambda = 0.50$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.29,.71,.88)	(.56,.74,.80)	(.72,.74,.76)	(.45,.75,.83)	(.72,.74,.76)	(.72,.74,.76)	(.65,.74,.78)	(.72,.74,.76)	(.72,.74,.77)
N(0.25,0.25^2)	(.27,.73,.84)	(.71,.74,.76)	(.72,.74,.76)	(.63,.74,.77)	(.72,.74,.76)	(.72,.74,.76)	(.71,.74,.76)	(.72,.74,.76)	(.72,.74,.78)
STRI(-0.5,1)	(.15,.33,.80)	(.34,.66,.90)	(.30,.54,.79)	(.29,.64,.86)	(.46,.61,.88)	(.71,.74,.76)	(.28,.66,.87)	(.72,.74,.76)	(.73,.74,.76)
SM(0,1,5)	(.16,.60,.79)	(.21,.67,.85)	(.43,.72,.82)	(.38,.74,.83)	(.73,.74,.75)	(.73,.74,.75)	(.71,.74,.77)	(.72,.74,.76)	(.71,.74,.78)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.79)	(.20,.71,.84)	(.29,.29,.89)	(.58,.74,.80)	(.73,.74,.75)	(.42,.73,.82)	(.72,.74,.76)	(.72,.74,.75)
U(0,0.5)	(.49,.75,.80)	(.72,.73,.77)	(.72,.73,.77)	(.71,.74,.76)	(.71,.74,.76)	(.71,.75,.76)	(.71,.74,.77)	(.72,.74,.77)	(.72,.74,.78)
Distribution of initial assessments	$\lambda = 0.75$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.37,.56,.90)	(.71,.74,.76)	(.72,.74,.76)	(.66,.73,.80)	(.72,.74,.76)	(.73,.74,.75)	(.72,.74,.76)	(.72,.74,.79)	(.72,.74,.75)
N(0.25,0.25^2)	(.33,.72,.83)	(.71,.74,.76)	(.71,.74,.76)	(.71,.73,.77)	(.72,.74,.76)	(.73,.74,.75)	(.72,.74,.76)	(.72,.74,.78)	(.67,.74,.78)
STRI(-0.5,1)	(.15,.33,.80)	(.32,.72,.87)	(.54,.73,.80)	(.24,.68,.82)	(.71,.74,.76)	(.72,.74,.76)	(.55,.74,.84)	(.72,.74,.76)	(.73,.74,.75)
SM(0,1,5)	(.26,.60,.77)	(.29,.71,.87)	(.41,.72,.83)	(.61,.74,.78)	(.72,.74,.76)	(.73,.74,.75)	(.72,.74,.76)	(.72,.74,.76)	(.70,.74,.79)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.17,.43,.81)	(.27,.65,.88)	(.73,.74,.76)	(.71,.74,.76)	(.70,.74,.77)	(.69,.74,.78)	(.72,.74,.75)
U(0,0.5)	(.33,.74,.84)	(.72,.75,.77)	(.72,.75,.77)	(.72,.74,.77)	(.72,.75,.77)	(.73,.74,.76)	(.71,.75,.76)	(.71,.74,.75)	(.68,.74,.78)
Distribution of initial assessments	$\lambda = 1.00$								
	h=0			h=6			h=12		
	n=25	n=50	n=75	n=25	n=50	n=75	n=25	n=50	n=75
U(-1,1)	(.25,.73,.89)	(.71,.74,.78)	(.71,.74,.78)	(.69,.74,.77)	(.73,.74,.76)	(.74,.74,.76)	(.70,.74,.77)	(.73,.74,.79)	(.73,.74,.79)
N(0.25,0.25^2)	(.22,.75,.88)	(.71,.73,.77)	(.71,.75,.77)	(.72,.74,.77)	(.73,.74,.76)	(.74,.74,.76)	(.71,.74,.79)	(.72,.74,.78)	(.71,.74,.79)
STRI(-0.5,1)	(.15,.33,.80)	(.32,.73,.86)	(.72,.73,.76)	(.27,.75,.86)	(.72,.74,.76)	(.72,.74,.77)	(.70,.74,.77)	(.68,.74,.78)	(.70,.74,.77)
SM(0,1,5)	(.24,.60,.78)	(.35,.71,.85)	(.54,.74,.83)	(.71,.74,.78)	(.73,.74,.76)	(.73,.74,.75)	(.68,.74,.78)	(.68,.74,.77)	(.68,.74,.75)
U(0.5,1)	(.21,.25,.81)	(.14,.54,.77)	(.15,.66,.82)	(.36,.72,.89)	(.73,.74,.76)	(.71,.74,.77)	(.70,.74,.76)	(.68,.74,.78)	(.68,.74,.78)
U(0,0.5)	(.31,.74,.82)	(.71,.74,.77)	(.71,.74,.77)	(.71,.74,.76)	(.72,.74,.76)	(.73,.74,.76)	(.68,.74,.79)	(.71,.74,.79)	(.68,.74,.79)

Table 2: Smirnov test statistics of the equality of the observed and the simulated empirical distributions of the agents' choices for the payoff assessment model with Parzen similarity

Period	G(2.44)	G(3.85)	Period	G(2.44)	G(3.85)	Period	G(2.44)	G(3.85)
1	0.18	0.18	26	0.22	0.16	51	0.29	0.25
2	0.12	0.25	27	0.27	0.16	52	0.24	0.23
3	0.14	0.22	28	0.26	0.12	53	0.28	0.25
4	0.20	0.18	29	0.21	0.12	54	0.23	0.23
5	0.33	0.22	30	0.31	0.13	55	0.23	0.25
6	0.30	0.31	31	0.30	0.23	56	0.27	0.30
7	0.28	0.26	32	0.32	0.15	57	0.27	0.30
8	0.16	0.14	33	0.26	0.15	58	0.26	0.30
9	0.23	0.30	34	0.31	0.13	59	0.24	0.30
10	0.17	0.31	35	0.31	0.15	60	0.20	0.23
11	0.30	0.24	36	0.31	0.15	61	0.26	0.30
12	0.22	0.10	37	0.30	0.13	62	0.25	0.30
13	0.19	0.20	38	0.31	0.15	63	0.25	0.30
14	0.22	0.13	39	0.31	0.20	64	0.25	0.30
15	0.27	0.15	40	0.29	0.20	65	0.24	0.30
16	0.24	0.10	41	0.28	0.20	66	0.25	0.30
17	0.19	0.12	42	0.28	0.20	67	0.19	0.35
18	0.20	0.12	43	0.30	0.20	68	0.24	0.30
19	0.22	0.13	44	0.31	0.25	69	0.23	0.30
20	0.23	0.13	45	0.30	0.25	70	0.24	0.30
21	0.20	0.18	46	0.30	0.15	71	0.25	0.30
22	0.15	0.24	47	0.32	0.23	72	0.25	0.30
23	0.15	0.14	48	0.26	0.20	73	0.25	0.30
24	0.23	0.14	49	0.26	0.25	74	0.26	0.28
25	0.23	0.12	50	0.29	0.25	75	0.26	0.30

Note: The Smirnov test statistic, better known as the two-way Kolmogorov-Smirnov statistic (Berry and Lindgren (1996), p.535), is based on the maximum vertical distance between the empirical distributions of the observed and the simulated choices. That is, $\max_x |F_m(x) - G_n(x)|$, where m and n are the number of simulated and observed choices and F_m and G_n are the empirical cumulative distribution functions of the simulated and observed choices, respectively. The 5 percent and 1 percent critical values for this test are 0.31 and 0.37 respectively.

Figure 1.a: Observed median choices of the four cohorts playing G(2.44)
(the interior equilibrium of G(2.44) is 0.59)

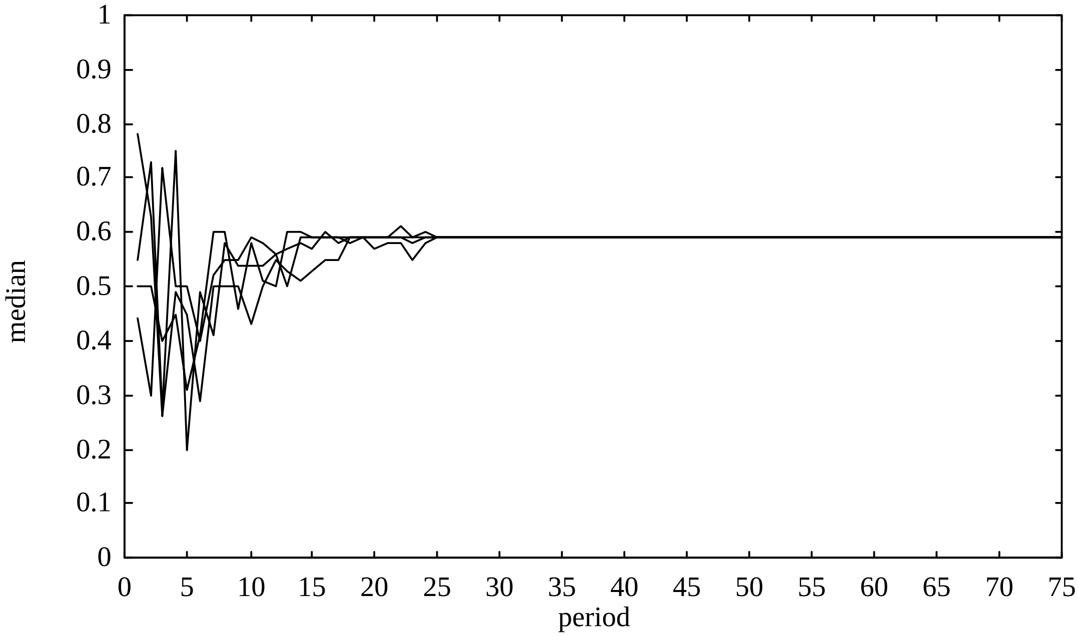


Figure 1.b: Observed medians of the four cohorts playing G(3.85)
(the interior equilibrium of G(3.85) is 0.74)

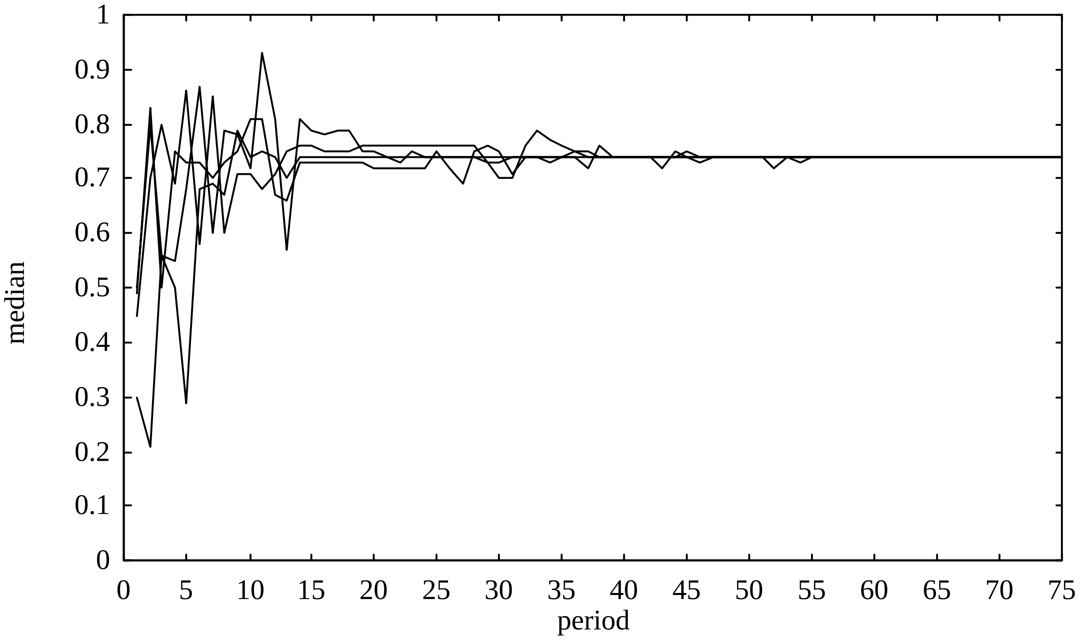


Figure 2.a: Observed choices for the four cohorts in $G(2.44)$

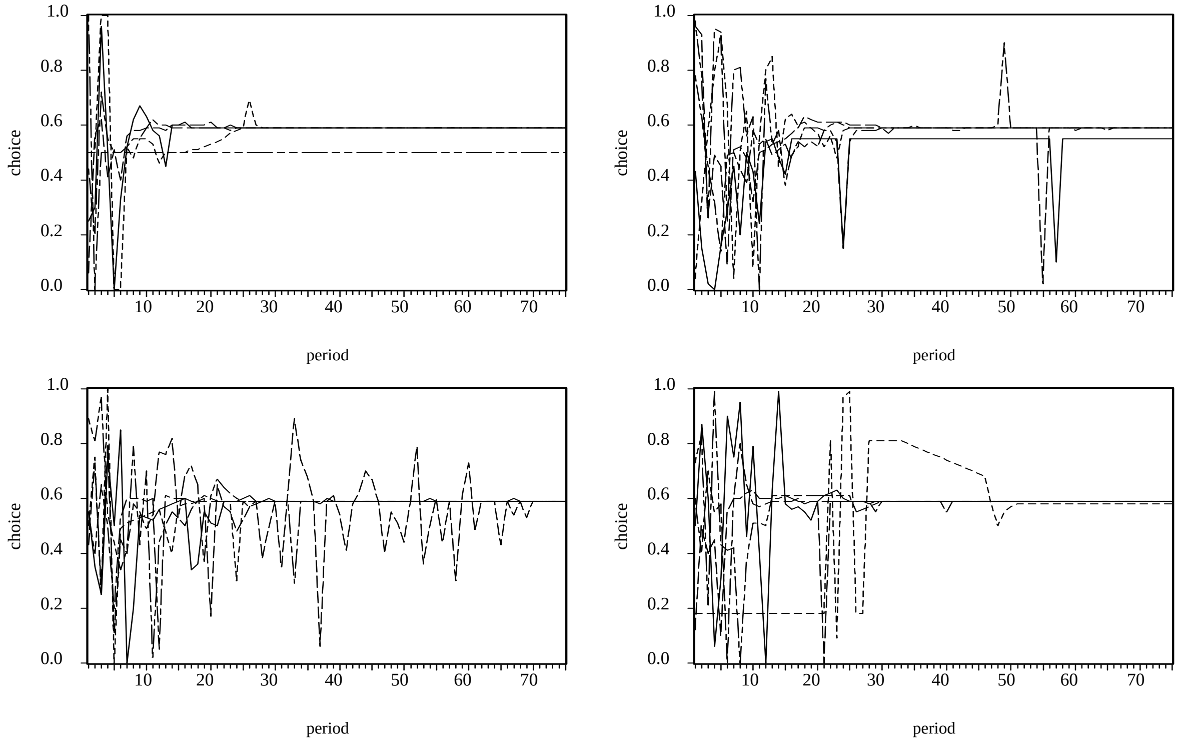


Figure 2.b: Observed choices for the four cohorts in $G(3.85)$

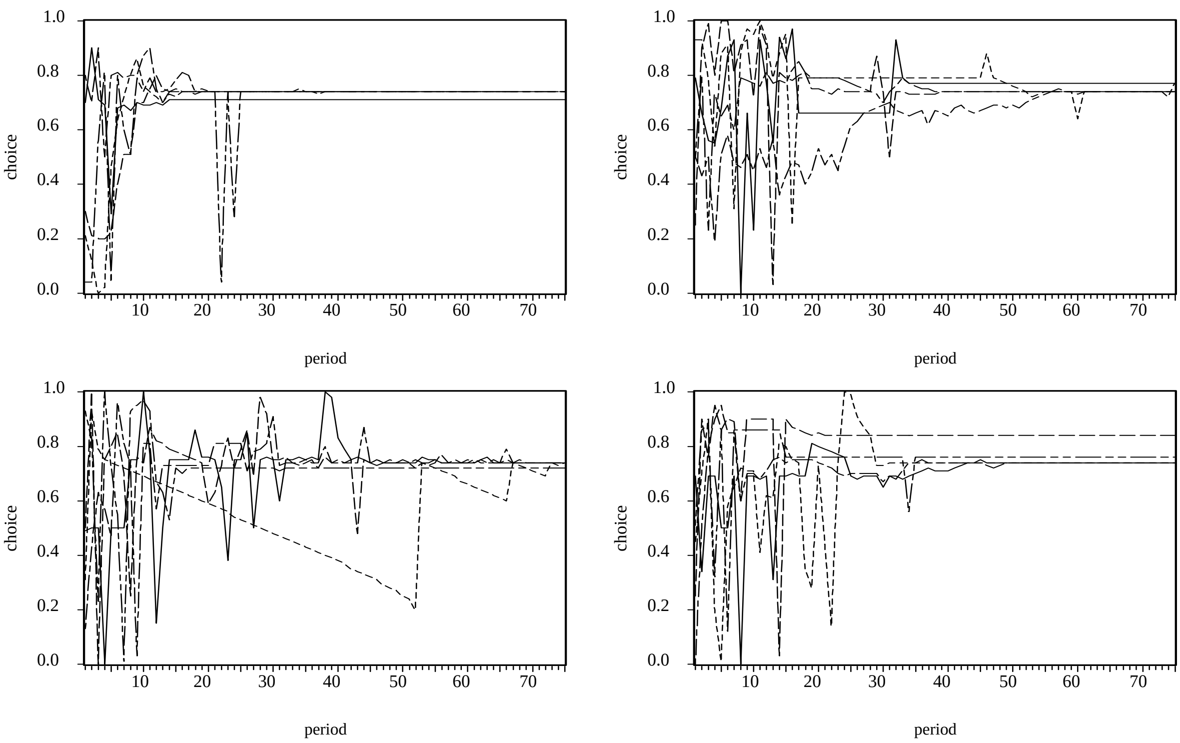


Figure 3: Histograms of final period choices for agents with U(-1,1) initial assessments and Parzen similarity function

$G(2.44), \lambda = 0.75, h = 6$

$G(2.44), \lambda = 1.00, h = 6$

$G(2.44), \lambda = 0.75, h = 12$

$G(2.44), \lambda = 1.00, h = 12$

$G(3.85), \lambda = 0.75, h = 6$

$G(3.85), \lambda = 1.00, h = 6$

$G(3.85), \lambda = 0.75, h = 12$

$G(3.85), \lambda = 1.00, h = 12$

Figure 4.a: A typical simulated path for the median of 5 players in $G(2.44)$
$(u(0) = U(-1, 1), h = 6, \lambda = 0.75)$

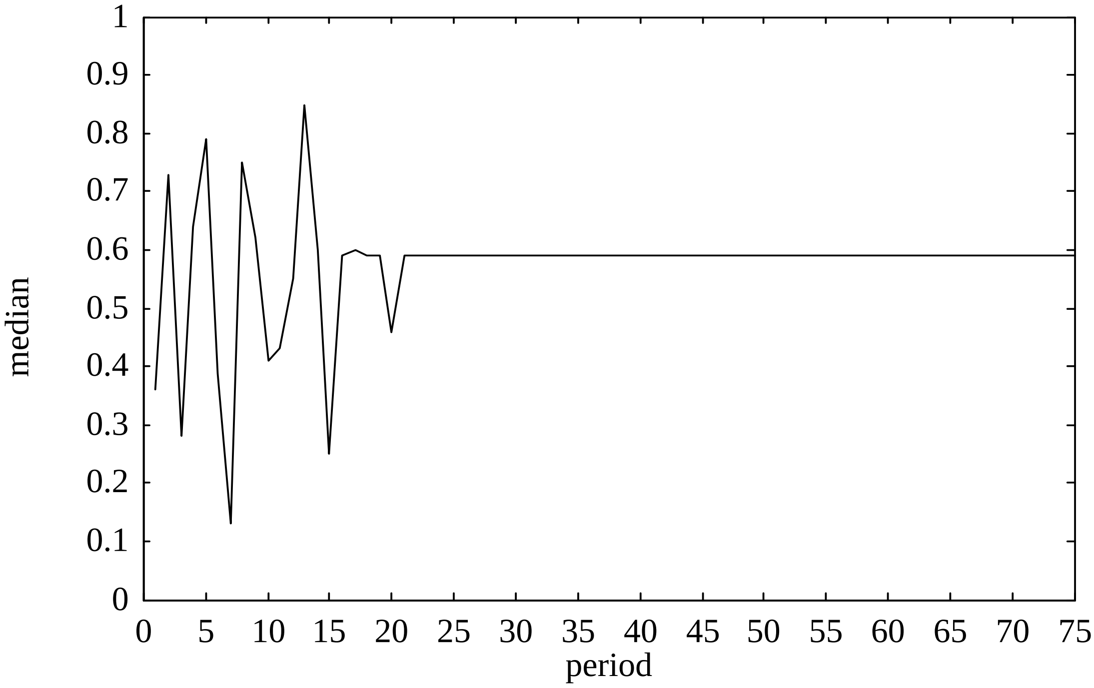


Figure 4.b: A typical simulated path for the median of 5 players in $G(3.85)$
$(u(0) = U(-1, 1), h = 6, \lambda = 0.75)$

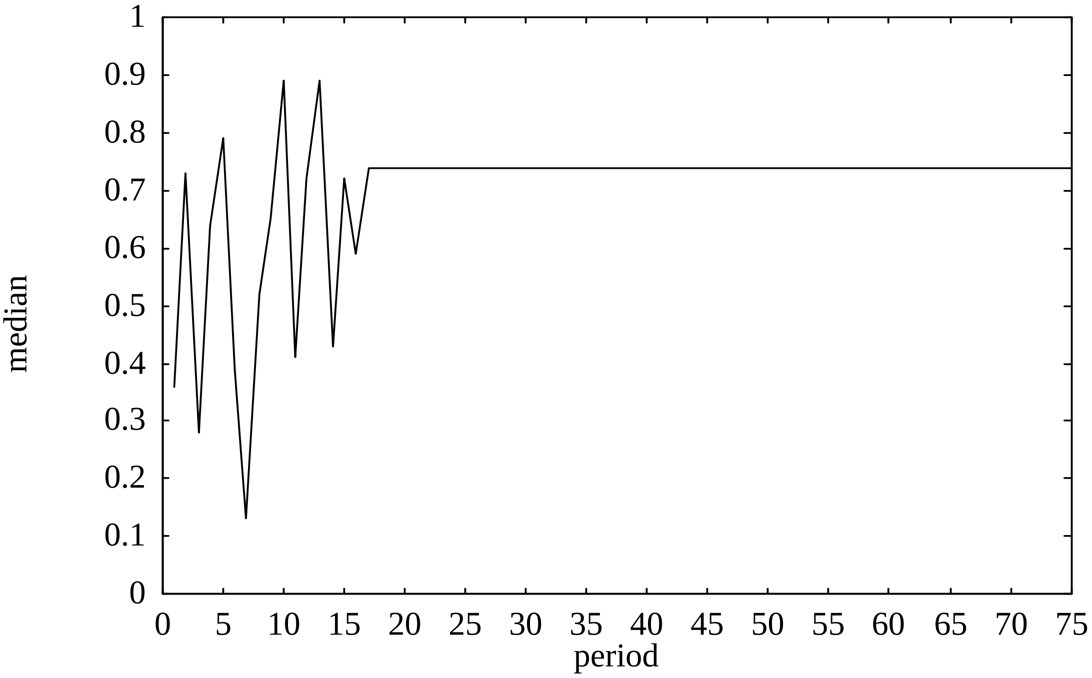


Figure 5: The objective function

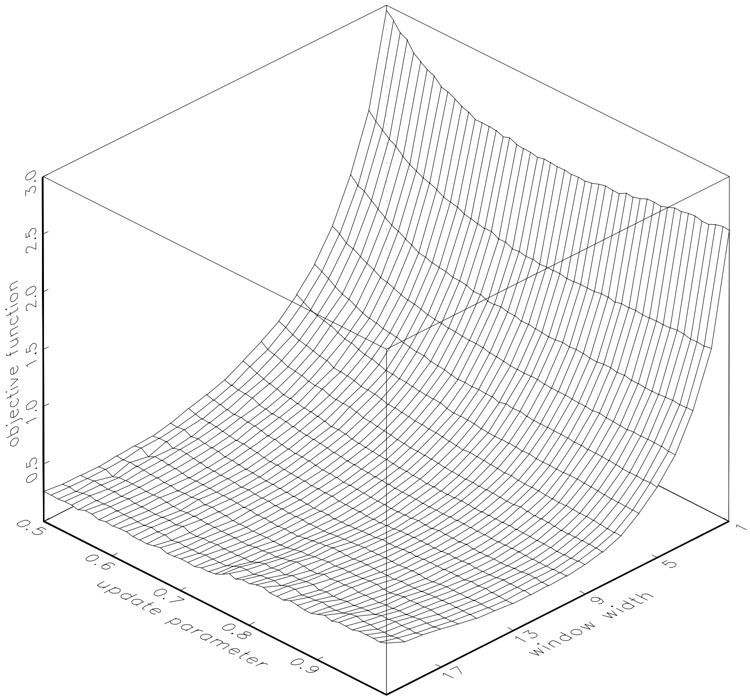


Figure 6.a: Histogram of the median choices in period 75 and 750 in 100 simulations of G(2.44) played by Cross reinforcement learners with Parzen similarity

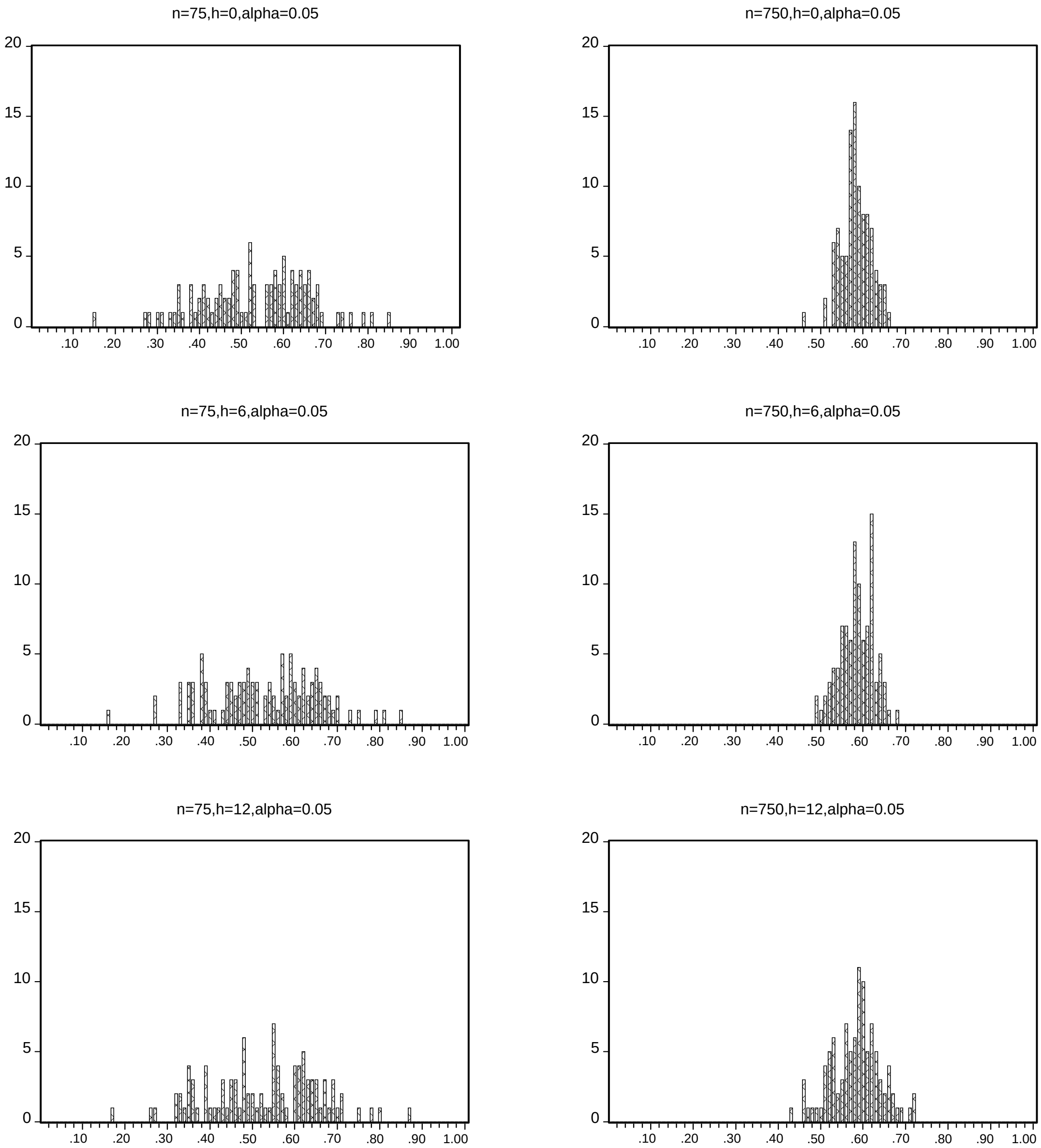


Figure 6.b: Histogram of the median choices in periods 75 and 750 in 100 simulations of G(3.85) played by Cross reinforcement learners with Parzen similarity

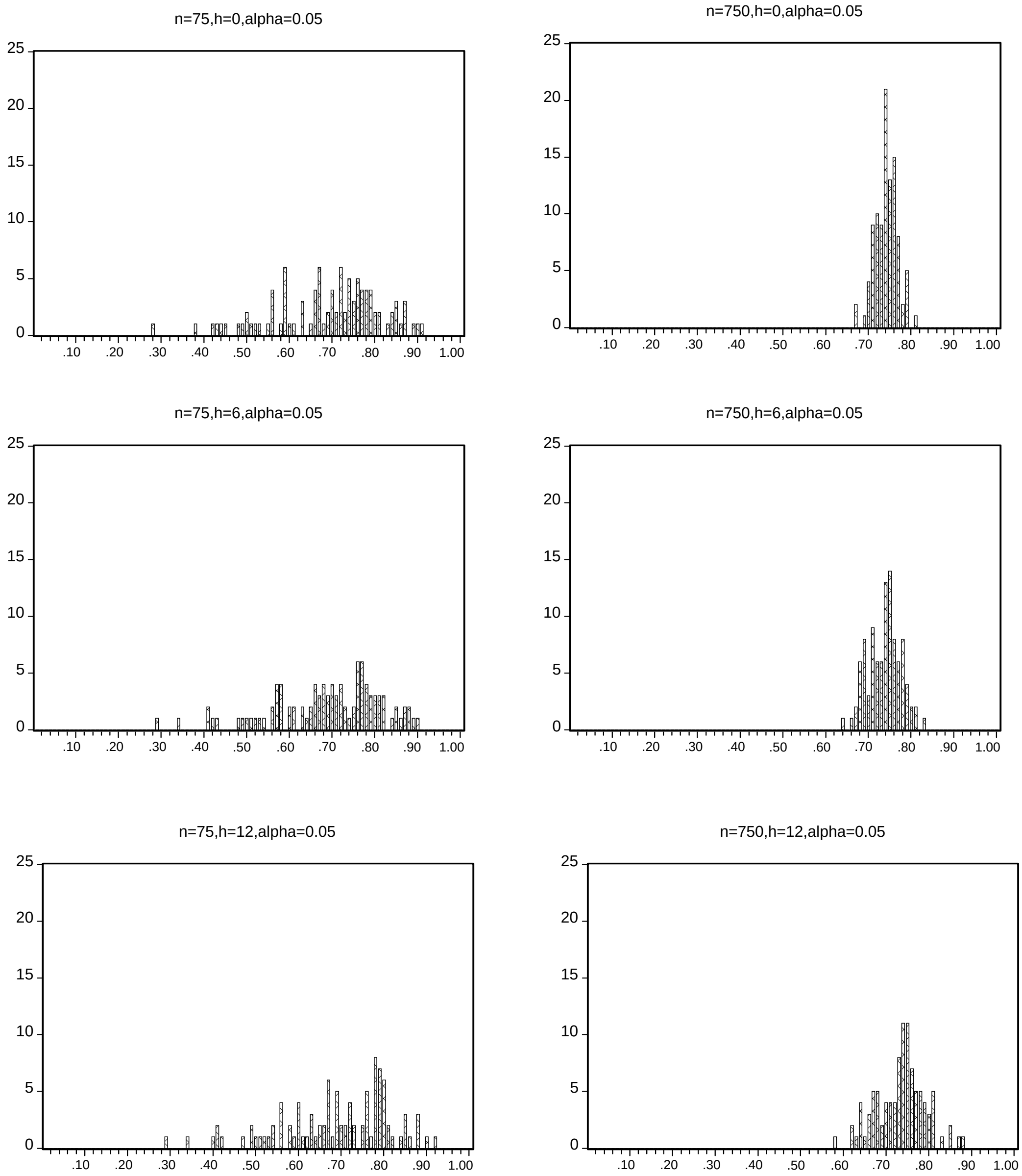


Figure 6.c: Histogram of the median choices in period 75 and 750 in 100 simulations of G(2.44) played by Cross reinforcement learners with Parzen similarity and declining learning parameter

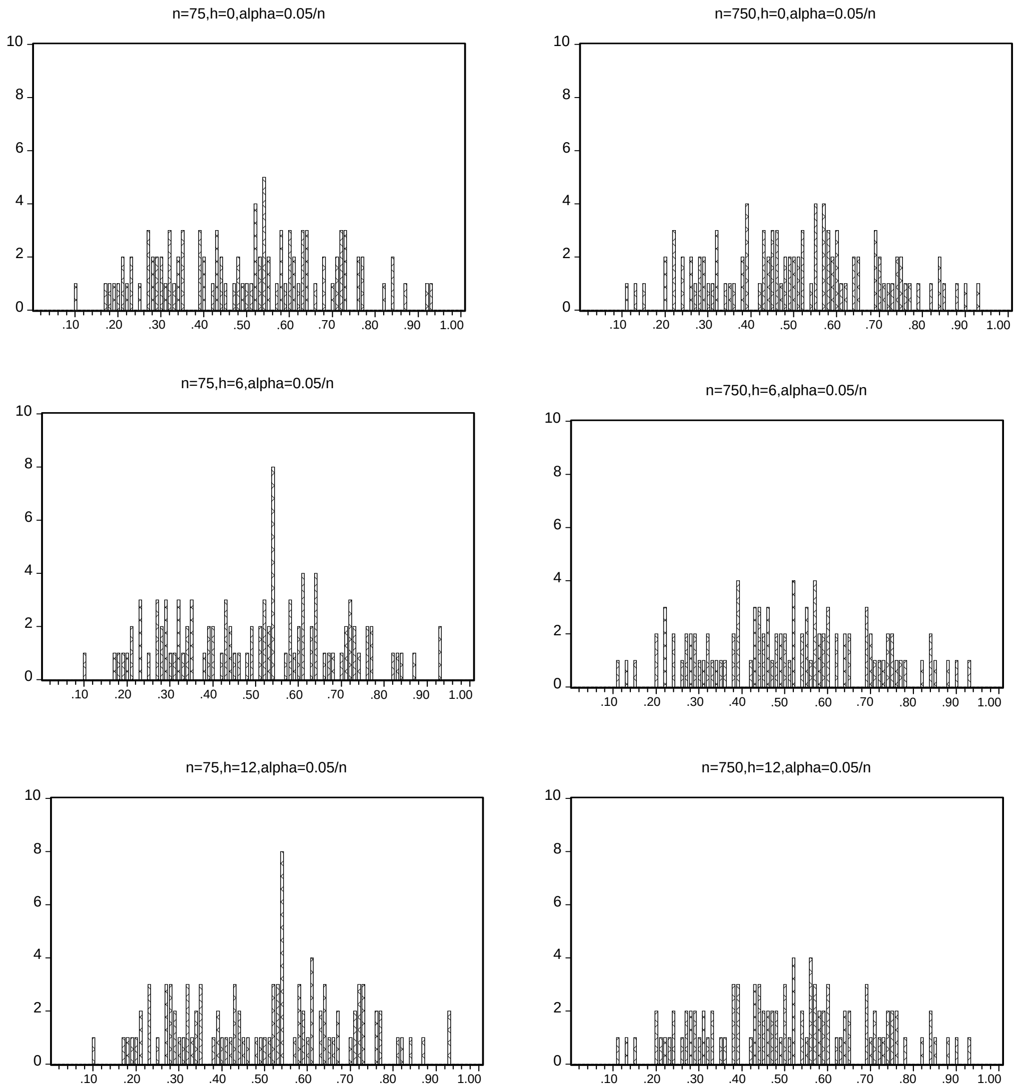


Figure 6.d: Histogram of the median choices in periods 75 and 750 in 100 simulations of G(3.85) played by Cross reinforcement learners with Parzen similarity and declining learning parameter