DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

A BAYESIAN SEQUENTIAL EXPERIMENTAL STUDY OF LEARNING IN GAMES

Mahmoud A. El-Gamal

Richard D. McKelvey

Thomas R. Palfrey

SOCIAL SCIENCE WORKING PAPER 757

February 1991
Revised August 1992

A Bayesian Sequential Experimental Study of Learning in Games

Mahmoud A. El-Gamal, Richard D. McKelvey and Thomas R. Palfrey

Abstract

We apply a sequential Bayesian sampling procedure to study two models of learning in repeated games. The first model is that individuals learn only about an opponent when they play her/him repeatedly, but do not update from their experience with that opponent when they move on to play the same game with other opponents. We label this the non-sequential model. The second model is that individuals use Bayesian updating to learn about population parameters from each of their opponents, as well as learning about the idiosyncrasies of that particular opponent. We call that the sequential model.

We sequentially sample observations on the behavior of experimental subjects in the so called 'centipede game'. This game has the property of allowing for a trade-off between competition and cooperation, which is of interest in many economic situations. At each point in time, the 'state' of our dynamic problem consists of our beliefs about the two models, and beliefs about the nuisance parameters of the two models. Our 'choice' set is to sample or not to sample one more data point, and if we should not sample, which of the models to select. After 19 matches (4 subjects per match), we stop and reject the non-sequential model in favor of the sequential model.

Keywords: Preposterior Analysis, Sequential Sampling, Bayesian Learning, Experimental Design, Game Theory.

A Bayesian Sequential Experimental Study of Learning in Games

Mahmoud A. El-Gamal, Richard D. McKelvey and Thomas R. Palfrey *

1 Introduction

This paper applies a sequential Bayesian approach to the design and analysis of experiments to test , statistically, between competing theoretical models. The basic idea is simple, and its application has been advocated by statisticians for years. Yet, to our knowledge, this approach to the design and analysis of controlled laboratory experiments in economics has not ever been followed. We apply this approach to compare and evaluate two competing equilibrium models about learning in games, which differ in basic assumptions about how much players learn about unknown population parameters of the game they are playing.

We begin by assessing our priors over two competing models (and the specific parameters of each model). After each observation, we update our beliefs about the parameters of the competing models and reevaluate the statistical evidence for one against the other. We continue gathering evidence until our priors are sufficiently concentrated so that we judge the value of an additional data point is not worth the cost of collecting it.

In the specific class of games we are studying, there are two important sources of learning for the players. The first kind is associated with a phenomenon known to

* We acknowledge the financial support from NSF grant #SES9011828 to the California Institute of Technology. We also acknowledge the able research assistance of Mark Fey, Lynell Jackson and Jeffrey Prisbrey in setting up the experiments, recruiting subjects and running the experiments. We acknowledge the help of the Jet Propulsion Laboratory and its staff members for giving us access to their Cray XMP/18, and subsequently their Cray YMP2E/116. We thank Gary Lorden for a valuable discussion. We also thank, without implicating, two editors and one associate editor of *JASA* and three referees for their very careful and useful comments on earlier versions.

economists and game theorists as "reputation building". Basically, it means that if one plays a specific opponent over a sequence of moves, then one can use observations about the opponent's early play to make inferences about how this opponent is likely to play in subsequent moves. The opportunity for reputation building of this sort plays an important role in many games of interest to social scientists, and has been a topic of much recent theoretical research.

The second kind of learning is what we call "population learning". In many economic experimental designs, especially those used to study behavior in strategic situations, each subject plays similar games against a sequence of opponents. Thus, players have an opportunity to learn about population parameters that are relevant to expectations about how a randomly drawn opponent is likely to act. This has implications about how players will play against early opponents compared to how they may play against later opponents. For example, they may adjust their strategies against early opponents in order to gather information about population parameters, which provides valuable statistical information about future opponents. We call this "strategic learning", since players are altering their strategy explicitly to affect how much they learn. This type of learning, which is familiar in search theory, has been studied by Fudenberg and Kreps (1988) in the context of repeated games.

We compare two models of learning. The first model is that individuals learn only about an opponent when they play her/him repeatedly, but do not update (about population parameters) from their experience with that opponent when they move on to play the same game with other opponents. We label this the non-sequential model. The second model is that individuals use Bayesian updating to learn about population parameters from each of their opponents, as well as learning about the idiosyncrasies of that particular opponent. We call that the sequential model.

Our approach to studying learning over time is to treat an experimental session as one large game, instead of a collection of independent observations. Therefore, our models explicitly take into account the fact that subjects will play multiple times against multiple opponents. A Bayesian Nash equilibrium to this larger game incorporates any learning that arises from the above sources.

The purpose of this paper is twofold. First, we wish to find the solution to the full game, and to find whether the full Bayesian Nash equilibrium makes predictions that are in line with the observed regularities in the experimental data. Second, we wish to test whether such a model is more or less likely than a non-sequential model given the observed data.

The paper is organized as follows: In section 2, we introduce the game, and lay out the particular parametrization of our models. In section 3, we describe the type of equilibrium we wish to find, and discuss the numerical methods of solving for the full Bayesian Nash equilibrium of the game, as well as the Bayesian Nash equilibrium of the competing non-sequential model. In section 4, we discuss the experimental design that we implement and the results we obtain. In section 5, we discuss the robustness of our results to various modeling assumptins, and conclude the paper.

2 The Game, the Experimental Design, and Our Model

For a variety of reasons, the *centipede game* has received a great deal of attention in the last few years. Binmore (1986) gave it its name and discussed several of its interesting connections with the foundations of game theory. Rosenthal (1982) first studied it and observed its application to industrial organization. Aumann (1988), Megiddo (1986), Reny (1985), McKelvey and Palfrey (1992) and others have also investigated it. The related problems of the chain-store paradox and the repeated prisoners' dilemma have extensive literatures that are too long to list here.

The game has two players, whom we call the red and blue players. In the 3-move version considered here there is a pile of money on the table. The red player gets to move first, and s/he gets to take 80% of the pile ($4.00) and leave the blue player 20% ($1.00), or s/he can pass, in which case the pile doubles and it is the blue player's turn to move. Now the blue player can take 80% of the current pile ($8.00) leaving red with $2.00 or pass, in which case the pile doubles again, and it is the red player's turn. The red player can take $16.00 and leave the blue player $4.00, or pass again. If red passes that second time that s/he gets to choose, the piles double one last time and blue gets \$32.00, leaving red with \$8.00 and the game ends.

By standard game theoretic reasoning, if it is common knowledge among the players that they are both selfish expected utility maximizers then a simple backward induction argument will show that the red player should take 80% of the original pile on the very first move and the game should end there. On the other hand, it has been noted by several researchers (see e.g. Binmore (1986), Reny (1985), Rosenthal (1982)) that there are difficulties with the use of backward induction in the centipede game, since observance of a pass gives information about a player's rationality, and hence might serve as a signal of her/his future behavior. Further, in the presence of the slightest lack of common knowledge of rationality, it is not necessarily any longer optimal for even the rational players to take (see e.g. Aumann (1988)). If one assumes that there is some small likelihood that the opponent misunderstood the game, or has altruistic preferences, then an argument similar to that which Kreps et al. (1982) used to explain cooperation in the finitely repeated prisoners' dilemma can be used to show that it may then be optimal even for a rational player to pass on early moves in the centipede game.

In addition to the theoretical arguments suggesting that one should expect some players to pass in the centipede game, McKelvey and Palfrey (1992) present some experimental evidence showing that in fact only a small proportion of players take on the first round. McKelvey and Palfrey find some support for the explanation based on lack of common knowledge of rationality. Moreover, in the above mentioned experiments, there is evidence of regular patterns of changes in subject behavior over time.

One other point is in order before we move on to the careful statement of the game and our modeling of it. Many game theoretic models make predictions that are too precise, in the sense that their error structures are degenerate. This is the case, for example, when Bayesian Nash equilibrium behavior predicts pure strategies. In laboratory data, this means that such models assign zero likelihood to many observed outcomes. This undermines the logic of rigorous statistical analysis in the context of structural models of this genre. We call this ***the zero likelihood problem.***

In order for a statistical analysis to be possible, it is necessary to either work with a model that admits all outcomes in the range of the data (e.g. Boylan and El-Gamal (1992)), or discard certain data points that the model predicts will never happen (e.g. Camerer and Weigelt (1988)). Since we would like to avoid such an ad hoc use of the data, we opt for the former approach.

We consider a four person game played in two stages. Players 1 and 2 are called red players, and players 3 and 4 are called blue players. The four players each play a sequence of two games as follows: At time 1, player 1 is matched with 3, and player 2 is matched with 4. Players 1 and 3 play the 3-move centipede game described above. Simultaneously, players 2 and 4 play the same game. At the conclusion of the first game, each player observes the outcome and payoff of the game that they are involved in but do not observe the outcome of the other game. Then at time 2, the process is repeated, but this time player 1 is matched with player 4 and player 2 is matched with player 3.

To an economist, it might seem strange at first to have subjects play the game only twice. Indeed, it is common practice in economic experiments to discard the first few "noisy" observations and analyze only data from later games. Since we are interested in "learning", however, it is precisely those early "noisy" data points about which we care. Indeed, later data points will not be of much use as most of the learning would have already taken place.

In modeling the game, we let nature take a move at the beginning of the game to decide on the rationality of each of our agents. We model nature's move as a random draw which makes any particular individual an altruist with probability q. An altruistic individual is assumed to pass all the time. The agents each have a subjective belief about the value of q, and we assume that each agent i has the belief that q is distributed uniformly on the interval $[0, \delta_i]$. Moreover, even though agents may have different values δ_i, they each believe that everyone else shares their own value δ_i. We then model the δ_i's to be drawn uniformly over the interval $[0, \epsilon]$, where ϵ is again a parameter of nature. There is also a common knowledge trembling probability α for both rationals and altruists. A tremble is a deviation of the player from the (optimal) equilibrium behavior due to temporary confusion, or any other source of making errors. In our model, if a player trembles at any move, s/he flips a fair coin to determine her/his move.

To avoid confusion about the various parameters, let us summarize the main points of the model:

- The probability of any person being an altruist is q, and q is assumed to be an objective parameter of nature that is unknown to us as well as to the players.

- Agent i believes that $q \sim U[0, \delta_i]$, and believes that $\delta_j = \delta_i$, $\forall j$. We assume that the actual δ_i's are in fact different for different players. This heterogeneity assumption embodies the notion that depending on the backgrounds of the subjects, they develop beliefs about the proportion of altruists in each sub-population. Players assume that everyone else in the experiment has similar beliefs about the sub-population from which the experimental subjects were sampled. The assumption of uniform prior is made for simplicity of computations.

- The actual δ_i's are distributed $U[0, \epsilon]$, and ϵ is assumed to be an objective parameter of nature which determines the distribution of backgrounds (and hence of δ_i's) in our pool of experimental subjects. Again the uniformity assumption is made for computational ease. The sensitivity analyses in section 5 suggest that our results are reasonably robust to that modeling assumption.

- Both rationals and altruists tremble with common knowledge probability α. This is assumed to be an objective parameter of nature that depends on the game and the subject pool from which we draw. Since our subject pool consists only of Caltech students, we assume that the complexity of the game immediately reveals to all of them the probability of "errors" being made by themselves as well as other players. We model those "errors" as complete confusion resulting in their flipping a fair coin to determine their move. This parameter is thus assumed to be common knowledge among the players, but unknown to us.

Each of our subjects will play the game with two players. Our two models of how subjects play the game are:

The Non-sequential Model: The subjects maintain their initial beliefs $q \sim U[0, \delta_i]$ and do not update from the first player's actions about the true value of q.

The Sequential Model: The subjects use Bayes's rule to update their beliefs about q between games.

Both models have three nuisance parameters from our point of view. These parameters are q, ϵ, and α. Our posteriors about those parameters are of interest in themselves if we maintain that we have the correct model of nature's and the subjects' behavior in this game. When we finish our sampling procedure and accept one of the two models, we shall indicate the values of q, ϵ, and α at which our posteriors are concentrated.

The experimental sessions were conducted by recruiting Caltech undergraduate students and informing them that they would participate in an decision making experiment where they would get to collect and keep what they earned in cash. Each experimental session required an even number of subjects who reported to the Caltech Laboratory for Experimental Economics and Political Science. The data used in this paper consist of 5 sessions with 16,16,16,12, and 16 participants, respectively. No participants were allowed in more than one session. When the participants arrived, they were seated at computer terminals with partitions that inhibit communication. They were also instructed that they should not communicate with any of the other participants at any point during the session. The participants drew an envelope at random, which informed them whether they were 'red' or 'blue'. They were told that they were either red or blue for the entire duration of the experimental session. Clear instructions were read aloud to all the participants. The exact rules of the three move centipede game were explained, and subjects were informed that they were playing for ***real money***, and would be paid exactly according to the rules, in cash, immediately following the session. The matching rules were also carefully explained, namely that they would play two games with different subjects of the other color. Example outcomes were illustrated and subjects were then given a quiz to assure us of their understanding of the structure of the experiment. At the end of the session, participants were paid their cash earnings in private in a separate room.

3 The Solution

For each value of the parameter α (which is assumed to be known to all players), and value of δ_i, a player's ***strategy*** is a vector of probabilities for passing at each move of the game. We first symmetrize the game by introducing an artificial chance move at the beginning of the game making each player red with probability 1/2 and blue with probability 1/2. This means that each player before the match begins must define the vector of probabilities of passing for all moves (red and blue) given her/his belief δ_i, and known probability of tremble α. Note that we only need to solve for the strategy of the rational players at the nodes of the game where they do not tremble, since altruists always pass, and all players flip fair coins when they tremble. Since each player may be red or blue, and each player plays two games with different opponents, the strategy of a player (who does not tremble) with initial belief δ and when the common knowledge tremble probability is α is represented by a ten dimensional vector $p(\delta,\alpha) \in [0,1]^{10}$. Where

- $p_1(\delta,\alpha)$= Probability that a rational red player passes in the first move of the first game.
- $p_2(\delta,\alpha)$= Probability that a rational blue player passes in the second move of the first game.
- For m=3,5,7,9, $p_m(\delta,\alpha)$= Probability that a rational red player passes in the first move of the second game, when the first game has ended in red taking on the first move (denoted T), blue taking on the second move (denoted PT), red taking on the third move (denoted PPT), or red passing on the third move (denoted PPP), respectively.
- For m=4,6,8,10, $p_m(\delta,\alpha)$= Probability that a rational blue player passes in the second move of the second game, when the first game has ended with a T, PT, PPT, or PPP, respectively.

So far, we have defined the strategy of a player with belief $U[0,\delta]$, when the true probability of trembling is α. How does a player choose a strategy that is *optimal* in some sense? We look for a ***symmetric Bayesian Nash equilibrium*** of the game, defined as the strategy $p(\delta, \alpha)$ which maximizes the player's total expected payoff from both games if all other rational players use the same strategy $p(\delta, \alpha)$.

A ***Bayesian Nash Equilibrium*** (Harsanyi (1967-68)) is a Nash equilibrium to a game in which there is some incomplete information about parameters affecting the payoffs of the players (in our case the altruism of the players). In a Bayesian Nash equilibrium, it is assumed that there is a joint distribution of player types which is common knowledge to the players. Each player conditions on her/his own type, and uses Bayes's rule together with her/his beliefs about the strategy of the other player to determine the conditional distribution of types of the other player, and in equilibrium, must act optimally at each decision node conditional on her/his updated priors at that node. We can alternatively consider the game as one in which player strategies are type contingent actions, and then a Bayesian Nash equilibrium is just a simple Nash equilibrium to this game. This is the approach we take below in computing the equilibrium. When one takes this approach, the updated priors do not actually have to be computed explicitly to achieve the solution, but they are implicit in the solution since they can be computed from it.

NOTES:

- We do not formally include a choice for a red player on the last move. In equilibrium, rational red players always take on the last move of both games.
- Remember that we are assuming that all players agree on the value of α, and believe that all players agree on the value of δ. This player-specific common knowledge assumption is consistent with the original formulation of incomplete information games of Harsanyi (1967-68). Most subsequent applied work in game theory assumed beliefs to be objectively true, or to be derivable from a "common prior". Neither of those assumptions is necessary in modeling games of incomplete information.
- As stated above, we only look for a symmetric Bayesian Nash equilibrium of the symmetrized game. Notice, however, that by symmetrizing the game in the way in-

dicated above, every equilibrium of the original game (symmetric and asymmetric) is equivalent to some symmetric equilibrium of the symmetrized game (e.g. Selten (1983)).

For a given δ and a given vector of probabilities $p = \{p_1, ..., p_{10}\} \in [0,1]^{10}$ representing a player's strategy, let us define $v(p, p'; \delta, \alpha)$ to be the expected payoff to the player when s/he uses the vector p and all other rational players use the vector p'. Then a symmetric Bayesian Nash equilibrium is defined as a vector $p^\star$ for each $(\delta, \alpha) \in [0,1]^2$, such that $v(p^\star, p^\star; \delta, \alpha) \geq v(p, p^\star; \delta, \alpha)$, $\forall p \in [0,1]^{10}$. For every $(\delta, \alpha) \in [0,1]^2$, the equilibrium vector $p^\star$ minimizes the function

$$Q(p; \delta, \alpha) = \sum_{p' \in \{0,1\}^{10}} \left(\max[\{v(p', p; \delta, \alpha) - v(p, p; \delta, \alpha)\}, 0] \right)^2$$

The function $Q: [0,1]^{10} \times [0,1]^2 \to \Re_+$ is non-negative and a zero of Q is by definition a Bayesian Nash equilibrium of the game. Further, Q can be shown to be everywhere differentiable (see e.g. McKelvey (1990)). Hence, standard gradient algorithms can be used to find a solution (a symmetric Bayesian Nash equilibrium of the game).

Our numerical computation of the equilibrium consists of evaluating the value function $v(p', p; \delta, \alpha)$ and finding the fixed point as described above. For the minimization part of the algorithm, we used the Powell's method routine in Press et al. (1988). The rest of the program was written in C and debugged on a 386 personal computer and a Sun 386i, and then ported and vectorized on a Cray XMP/18, and then a Cray YMP2E/116. We were content with bringing the function $Q(.)$ above to less than 10^{-20} at 100 values of δ and 20 values of α (a total of 2000 grid points).

Similarly, we solve for the symmetric Bayesian Nash equilibrium of the non-sequential model which is much simpler since the two probabilities of red passing on the first move and blue passing on the second should be the same for the first and second games (regardless of the outcome of the first game). This is simply the definition of non-sequential behavior: the agent does not use the information from the first game to update her/his belief on q and choose a different set of probabilities. Hence, the non-sequential equilibrium is represented by a vector in $[0,1]^2$.

4 The Experimental Design

We begin with a prior about the relative validity of the two learning models which we represent by the prior odds ratio. We decided to start with an uninformed prior odds ratio of unity. We also begin with an uninformative prior on the three unknown (nuisance) parameters of nature q, ϵ, and α, which is given by the uniform prior density over $[0,1]^3$.

Now, we collect our data with matches of 4 players where players 1 and 2 are red, and players 3 and 4 are blue. After each match, we revise our beliefs about the three nuisance parameters under each of the two models, and also revise our posterior odds ratio. Given our prior on (q, ϵ, α), we compute the likelihood of the two models by integrating out the nuisance parameters with respect to that prior. For the Sequential Nash model, we start by computing for each player $i = 1, 2, 3, 4$ the likelihood of that player's moves if s/he is rational and has belief δ_i:

$$like_{sn}^{i,rat}(match; \delta_i, \alpha) = \prod_{\text{i's moves}} p_{sn}(move, \delta_i, \alpha).$$

where

$$p_{sn}(pass, \delta_i, \alpha) = (1-\alpha) \ \text{argmin}_p \ Q(p; \delta_i, \alpha) + \alpha/2,$$

and

$$p_{sn}(take, \delta_i, \alpha) = (1-\alpha)(1 - \text{argmin}_p \ Q(p; \delta_i, \alpha)) + \alpha/2.$$

Similarly, we compute the likelihood of the same player if s/he is an altruist and has belief δ_i:

$$like_{sn}^{i,alt}(match; \delta_i, \alpha) = \prod_{\text{i's moves}} p^{alt}(move).$$

where $p^{alt}(pass) = (1-\alpha) + \alpha/2$, and $p^{alt}(take) = \alpha/2$.

Now, the likelihood of player i's moves conditional on the proportion of altruists being q is

$$like_{sn}^{i}(match; \delta_i, \alpha, q) = (1-q) \cdot like_{sn}^{i,rat}(match; \delta_i, \alpha) + q \cdot like_{sn}^{i,alt}(match; \delta_i, \alpha)$$

Now, given our assumption that the δ_i's are distributed $U[0, \epsilon]$, and integrating out

(ϵ, q, α) with respect to our prior, the full likelihood of the match can be written as:

$$like_{sn}(match) = \int\limits_{[0,1]^3} l_{sn}(match; \epsilon, q, \alpha)\ prior(d\epsilon, dq, d\alpha)$$

where

$$l_{sn}(match; \epsilon, q, \alpha) = \prod_{i=1}^{4} \left(\int_0^{\epsilon} like^i_{sn}(match; \delta_i, \alpha, q)\ d\delta_i \right)$$

Similarly, we can compute $like_{nn}(match)$ in the same manner for the non-sequential model, replacing the probabilities by their non-sequential counterparts.

After observing the outcome of a match, we do two updates. First, we update the posterior odds ratio on the two models by multiplying by the likelihood ratio given the match outcome. We also update our beliefs about (ϵ, q, α) for each of the hypotheses via Bayes's rule. Our posterior belief under SN on the nuisance parameters becomes

$$posterior_{sn}(d\epsilon, dq, d\alpha) = \frac{l_{sn}(match; \epsilon, q, \alpha)\ prior_{sn}(d\epsilon, dq, d\alpha)}{\int\limits_{[0,1]^3} l_{sn}(match; \epsilon, q, \alpha)\ prior_{sn}(d\epsilon, dq, d\alpha)}$$

and similarly for the non-sequential Nash model, our posterior belief on the nuisance parameters becomes

$$posterior_{nn}(d\epsilon, dq, d\alpha) = \frac{l_{nn}(match; \epsilon, q, \alpha)\ prior_{nn}(d\epsilon, dq, d\alpha)}{\int\limits_{[0,1]^3} l_{nn}(match; \epsilon, q, \alpha)\ prior_{nn}(d\epsilon, dq, d\alpha)}$$

Now, we address the issue of the stopping rule. It is well known in the case of two densities (if we did not have the nuisance parameters) (e.g. see Chernoff (1972, Ch. 11,12), Širjaev (1973, Ch. IV.1), and Berger (1985, Ch.7)) that for (0-K) loss functions (where K is the loss if we stop sampling and select the wrong model), the sequential probability ratio test which stops when the odds ratio hits one of two boundaries is a Bayes procedure. This procedure also has the remarkable property (see any of the above references) that it stops in finite time with probability 1, and simultaneously minimizes the expected stopping time under both models among all stopping procedures with the same type I and type II error probabilities.

In the case where the two likelihood functions depend on nuisance parameters, in the tradition of Chernoff (1959) and Kiefer and Sacks (1963), one considers asymptotically (as the cost per data point goes to zero) optimal Bayes procedures. Those procedures (Kiefer and Sacks (1963), Lorden (1977)) stop when the odds ratio hits one of two boundaries that depend on the cost of sampling and some information numbers. The odds ratio is the ratio of the likelihoods of the two models, where each of the likelihoods is computed under each (ϵ, q, α) vector and then integrated with respect to our prior on those nuisance parameters.

Accordingly, we implemented the following version of the procedure: After each match, we compute the likelihood of each of the 256 possible data sets under each of the model, and calculate the Kullback-Liebler information numbers

$$\mu_{sn} = E_{sn}[\log(like_{sn}(match)/like_{nn}(match))]$$

and

$$\mu_{nn} = E_{nn}[\log(like_{sn}(match)/like_{nn}(match))].$$

By using the Wald approximation of probabilities of type I and type II error and the expected stopping time (Chernoff (1972, pp. 59-66), Berger (1985, pp. 485-499)), we then compute the boundaries A and B, $0 < A < 1 < B$, such that we stop and accept the sequential model if our posterior odds ratio exceeds B, and stop and accept the non-sequential model if our posterior odds ratio is smaller than A. Following the prescription in Berger (1985, p. 500), an approximately optimal choice of A and B involves minimizing the approximate risk function:

$$r(\pi, d^{A,B}) \approx \pi\left\{\frac{(1-A)}{(B-A)}K + \frac{c}{\mu_{nn}}\left[(\log A) + (\log\frac{B}{A})\frac{(1-A)}{(B-A)}\right]\right\}$$
$$+(1-\pi)\left\{\frac{A(B-1)}{(B-A)}K + \frac{c}{\mu_{sn}}\left[(\log A) + (\log\frac{B}{A})\frac{B(1-A)}{(B-A)}\right]\right\}$$

with respect to A and B. We set c (the cost per unit sampled) at 0.01 , K (the loss of selecting the wrong hypothesis) at 1; π denotes our prior belief on NN , and μ_{nn} and μ_{sn} are the information numbers calculated with our priors on (ϵ, q, α). Our choice of c and K does not directly reflect the actual dollar cost of running an extra game. It reflects our preferences against the selection of the wrong hypothesis by choosing the cost (in utility terms) to be $1/100^{th}$ of the loss (again in utility terms) of selecting the wrong hypothesis.

Since we chose $c = 0.01 << K = 1$, we use the further approximation (see previous references) for small c which yields $A = -c\pi/(\mu_{nn}K(1-\pi))$, and $B = \pi K\mu_{sn}/(c(1-\pi))$. Our stopping criteria are not necessarily optimal for a number of reasons: First, we use an approximation as $c \to 0$, and, second, we ignore the problem of overshooting the boundaries (Lorden (1977)). Berger did not prove that the formulae he recommended (Berger (1985, 7.83 and 7.84, p.499)) are optimal, but we use them since they seem to be sensible means for computing the stopping boundaries.

An alternative approach that we can follow, even though its optimality properties are unknown, is to recompute the boundaries at each point using our updated odds ratio π^t, and information numbers, μ^t_{sn} and μ^t_{nn}, calculated using the updated priors on the nuisance parameters. The latter procedure seems intuitively appealing since it recalculates the Bayes risk at each point using all the information to-date, and then decides whether or not to sample depending on that posterior Bayes risk. We include both the constant boundaries and the varying ones in Figure 1. It turns out in our case that both procedures yield exactly the same result, as seen in Figure 1.

RESULTS:

Table 1 shows the data we collected from the 19 matches in the 5 experimental sessions, and reports the posterior odds ratio for SN versus NN after each match. The posterior odds ratio is also plotted as the solid line in Figure 1. The stopping boundaries computed à la Berger (1985, p. 500) are plotted as the two flat dashed lines in Figure 1. If we follow the second procedure of recomputing the boundaries in each period using the current posterior odds ratio and posteriors on the nuisance parameters, we obtain the two changing boundaries shown by the gray lines in Figure 1. After 19 matches, we stop and accept the Sequential Nash model.

Match no.	1-3 Game	2-4 Game	1-4 Game	2-3 Game	odds ratio
Before observing data					1.00000
Match 1	PT	T	PPT	PT	0.94137
Match 2	PPT	PT	PT	T	1.03144
Match 3	PT	PT	T	PT	1.14797
Match 4	T	T	T	T	1.06996
Match 5	PT	PT	PT	T	1.12330
Match 6	T	T	T	T	0.89294
Match 7	PT	PPP	PPT	PPT	2.91379
Match 8	T	PPP	PPT	PPP	6.68167
Match 9	T	PPP	PPT	PPT	9.09750
Match 10	PT	PPT	PT	PPT	6.62331
Match 11	PT	PPP	PPT	PT	1.17326
Match 12	PT	PPT	PPP	PT	1.25507
Match 13	T	PT	T	PT	1.09545
Match 14	PT	T	T	T	1.62433
Match 15	PPP	T	PPP	T	2.60064
Match 16	T	PPP	PPT	PPP	5.19881
Match 17	PPT	PPT	PPT	PPT	4.37539
Match 18	T	PPT	T	PT	5.24771
Match 19	PPT	PPT	PPP	PPT	4.67080

Table 1: Evolution of odds ratio

REMARKS:

- We inherited the first 12 matches from an earlier version of the paper which had a different model of nature and player beliefs. Therefore, we could not pretend to be able to stop before 12 matches had passed.

- For purposes of anonymity, we wanted to have a large number of agents during an experimental session so that players will not know who is in their group of 4. As a result, we ran sessions typically with 16 subjects, with the exception of one session where we had 12 subjects. The way we order matches within an experimental session is arbitrary (according to subject number in the session). In Figure 1, we separate the matches between laboratory session by vertical lines. Since data were gathered in batches of matches, we could only have stopped sampling at such points. The lines at 4 and 8 are gray to indicate that the first twelve matches were data inherited from the previous version of the paper.

- We plot 10 times log of base 10 of the odds ratio and stopping barriers (i.e. measured in decibels) to make the graph more symmetric, and show the variations in the lower boundary.

- Our joint posteriors on (ϵ, q, α) are concentrated. The mode of the posterior are at $\epsilon = 0.25$, $q = 0.2$, and $\alpha = 0.15$ under SN; and $\epsilon = 0.2$, $q = 0.2$, and $\alpha = 0.2$ under NN. The means of our joint posterior are at $\epsilon = 0.21$, $q = 0.23$, and $\alpha = 0.17$ under SN; and $\epsilon = 0.25$, $q = 0.21$, and $\alpha = 0.17$ under NN. The standard deviations of our joint posterior are at 0.07 for ϵ, 0.09 for q, and 0.05 for α under SN; and at 0.09 for ϵ, 0.1 for q, and 0.05 for α under NN. Those parameters suggest about 15-20% of the time subjects tremble (are confused and decide to flip a coin), and that the proportion of altruists is approximately 20%, with subjects strongly underestimating that proportion.

- Notice that had we had our current model when we started collecting data, we would have stopped after the eighth match. Since after 12 matches our posterior odds ratio was back near 1, we continued sampling. Fortunately, the decision we make after 19 matches is the same that we would have made after 8 (accepting the sequential model).

5 Robustness and Limitations of Our Results:

The question may arise: how does our result of accepting the sequential model over the non-sequential one depend on the particular parametrization and distributional assumptions of the model? One can think of at least three levels at which one might consider the robustness of the results to the particular assumptions we make. These levels, going from the highest to the lowest are:

1. How robust are our results to our uniform priors we place on the nuisance parameters (q, ϵ, α)?

2. How robust are our results to the assumption that the true δ_i's are distributed uniformly over $[0, \epsilon]$ for some epsilon?

3. How robust are our results to the assumption that players formulate beliefs on q using a uniform density over $[0, \delta_i]$.

The easiest to answer is the first robustness question, and the answer is that our results are robust in that respect. As can be seen from the posterior standard deviations on the nuisance parameters, the likelihood function very strongly dominates the prior. For all reasonable priors (i.e. ones that do not concentrate mass in a small interval of high values for ϵ's, α's, or q's), our results would be unchanged.

For the second robustness question, one would ideally like to have a very large class of potential models (i.e. parametrizations of the distribution of the δ_i's) and compute the likelihoods under the sequential and non-sequential models for each of those parametrizations. This paper presented only two potential models (namely SN and NN). We subsequently conducted a sensitivity analysis, be recalculating all of the likelihoods of SN and NN (based on the entire data set) under four alternative specifications. For completeness, we include in the discussion below the model we used in this paper (Model 1). The models we use for our sensitivity analysis are:

Model 1: Let $\delta_i \sim U[0, \epsilon]$ and place prior on $\epsilon \sim U[0, 1]$.

Model 2: Let $\delta_i \sim U[0.1 - \epsilon, 0.1 + \epsilon]$ and place prior on $\epsilon \sim [0, 0.1]$.

Model 3: Let $\delta_i \sim U[\epsilon, 0.05]$ and put prior on $\epsilon \sim U[0, 0.05]$.

Model 4: Let $\delta_i \sim U[\epsilon, 0.3]$ and put prior on $\epsilon \sim U[0, 0.3]$.

Model 5: Let $\delta_i \sim U[\epsilon, 0.5]$ and put prior on $\epsilon \sim U[0, 0.5]$.

Recall that in Model 1, the mean of our posterior estimate of the average δ_i is approximately 0.125 ($\epsilon \approx 0.25$). Thus, we chose a range of alternative models that either force this posterior estimate to be lower (Model 3), or higher (Models 4 and 5). Model 2 forces the estimate of the average δ_i to be 0.1, but permits the variance of the δ_i's to be different. The resulting likelihoods for the various models are shown in Table 2 below. It turns out that the Models 2-5 all tend to nest themselves into Model 1 (by concentrating the posterior on the nuisance parameter ϵ at 0.1 for Model 2, and at 0 for models 3-5). They are constrained from achieving the same distribution $\delta_i \sim U[0, 0.25]$ of Model 1, and they all have smaller likelihoods than Model 1 under both SN and NN. Notice, moreover, that the models with the distribution of δ_i's farthest from Model 1, have the lowest likelihoods (e.g. Model 5 is about 100,000 times less likely than Model 1 under SN and 1000 times less likely under NN). All computations in this table were done with the marginal priors on q and α still being $U[0, 1]$. The results of the following table suggest that our model is reasonably robust in the sense that even if we incorporated these other models in the analysis, the SN version of Model 1 will still be the most likely. In fact, the only alternative model in which the NN version is more likely than the SN version (Model 5) is strongly rejected by ***all*** the other models.

TABLE 2					
SENSITIVITY ANALYSIS					
LIKELIHOODS OF VARIOUS MODELS					
Model	Dist. of δ_i	Prior on ϵ	Like. under SN	Like. under NN	odds ratio
1	U[0,ϵ]	U[0,1]	9.2463×10^{-43}	1.9796×10^{-43}	4.67
2	U[0.1-ϵ,0.1+ϵ]	U[0,0.1]	2.0858×10^{-43}	4.1367×10^{-44}	5.04
3	U[ϵ,0.05]	U[0,0.05]	4.4845×10^{-43}	2.6894×10^{-44}	16.68
4	U[ϵ,0.3]	U[0,0.3]	1.331×10^{-43}	6.0985×10^{-44}	2.18
5	U[ϵ,0.5]	U[0,0.5]	2.6728×10^{-48}	5.5817×10^{-46}	0.005

Now, the only remaining assumption is that regarding the way in which players formulate priors on q. Priors on q that entertain the possibility of very low q will result in a larger proportion of takes, and priors that put most of the mass on high values of q will result in more passing. Point mass distributions on q will result in no difference between the sequential and nonsequential models. An actual sensitivity analysis is not feasible since the computation of the equilibrium for any other parametrization will cost us hundreds more of supercomputer CPU hours. Finally, we note that the results reported here may be limited to the particular class of games that we are studying, and hence should not be interpreted as universal statements about the classes of models that we compare.

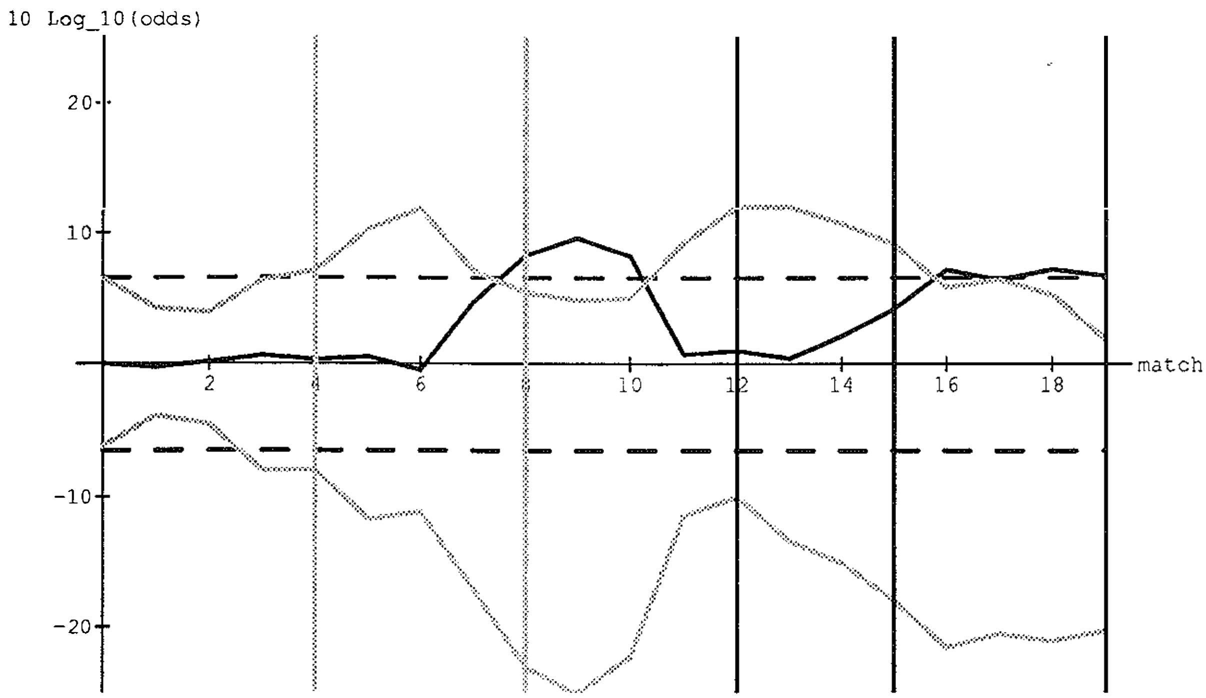

Figure 1: Time series of 10 $\log_{10}$ of odds ratio (solid line), and stopping boundaries (dashed lines, and gray lines)

References

Aumann, R. 1988. Preliminary notes on integrating irrationality into game theory. Mimeo.

Berger, J. 1985. *Statistical decicion theory and Bayesian analysis*. Springer-Verlag.

Binmore, K. 1986. Modeling rational players. Technical report, London School of Economics.

Boylan, R. and M. El-Gamal. 1992. Fictitious play: A statistical study of multiple economic experiments. *Games and Economic Behavior*. In Press.

Camerer, C. and K. Weigelt. 1988. Experimental tests of a sequential equilibrium reputation model. *Econometrica* 56:1–36.

Chernoff, H. 1959. Sequential design of experiments. *Annals of Mathematical Statistics* 30:755–770.

Chernoff, H. 1972. *Sequential analysis and optimal design*. Philadelphia: SIAM.

Fudenberg, D. and D. Kreps. 1988. A theory of learning, experimentation and equilibrium in games. Mimeo.

Harsanyi, J. C. 1967-68. Games of incomplete information played by Bayesian players (I, II, and III). *Management Science* 14:159–182, 320–334, 486–502.

Kiefer, J. and J. Sacks. 1963. Asymptotically optimum sequential inference and desgn. *Annals of Mathematical Statistics* 34:705–750.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson. 1982. Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory* 27:245–252.

Lorden, G. 1977. Nearly optimal sequential tests for finitely many parameter values. *Annals of Statistics* 5:1–21.

McKelvey, R. D. 1990. A Liapunov function for Nash equilbiria. Mimeo. California Institute of Technology.

McKelvey, R. D. and T. Palfrey. 1992. An experimental study of the centipede game. *Econometrica* 60:803–836.

Megiddo, N. 1986. Remarks on bounded rationality. IBM Research Report.

Press, W., B. Flannery, S. Teukolsky, and W. Vetterling. 1988. *Numerical recipes in C: The art of scientific computing.* Cambridge: Cambridge University Press.

Reny, P. 1985. Rationality, common knowledge, and the theory of games. Mimeo. Princeton University.

Rosenthal, R. 1982. Games of perfect information, predatory pricing, and the chain store paradox. *Journal of Economic Theory* 25:92–100.

Selten, R. 1983. Evolutionary stability in extensive two person games. *Mathematic Social Sciences* 5:269–363.

Širjaev, A. N. 1973. *Statistical sequential analysis.* Providence, Rhode Island: American Mathematical Society.