A hidden Markov model for the detection of pure and mixed strategy play in games*

Jason Shachat[†]

J. Todd Swarthout[‡]

Lijia Wei[§]

July 7, 2012

Abstract

We propose a statistical model to assess whether individuals strategically use mixed strategies in repeated games. We formulate a hidden Markov model in which the latent state space contains both pure and mixed strategies, and allows switching between these states. We apply the model to data from an experiment in which human subjects repeatedly play a normal form game against a computer that always follows its part of the unique mixed strategy Nash equilibrium profile. Estimated results show significant mixed strategy play and non-stationary dynamics. We also explore the ability of the model to forecast action choice.

[†]Wang Yanan Institute for Studies in Economics and MOE Key Laboratory in Econometrics, Xiamen University. *jason.shachat@gmail.com*

[‡]Department of Economics and Experimental Economics Center, Georgia State University. *swarthout@gsu.edu*

[§]Wang Yanan Institute for Studies in Economics and MOE Key Laboratory in Econometrics, Xiamen University. *ljwie.wise@gmail.com*

1

1   Introduction

Game theory and the Nash equilibrium solution concept are a key framework in the social sciences for modeling interactive behavior. The formulation of a normal form game consists of a set of players, a set of possible actions for each player, and a payoff function for each player that gives a real-valued payoff for any possible joint action profile – a list of actions consisting of one for each player. A Nash equilibrium is a joint action profile such that each player's assigned action results in at least as high a payoff to the player as any other possible action, assuming all other players choose their respective actions in the Nash equilibrium profile. If players are restricted to deterministically choose an action, then there are many games that don't have a Nash equilibrium, such as the childhood game of Rock, Scissors, Paper. Confronted with this problem, Von Neumann (1928) generalized a player's decision from choosing an action to choosing a probability distribution over his possible actions.[1] This choice of a probability distribution is called a "mixed" strategy, and a degenerate mixed strategy which chooses a particular action with probability one is called a "pure" strategy. The introduction of mixed strategies allows for existence of equilibrium across a broad class of games: from minimax solutions for zero-sum games (Von Neumann, 1928; Von Neumann and Morgenstern, 1944) to noncooperative equilibria for n-person games (Nash, 1951). While the role of mixed strategies in defining logically consistent solution concepts is not in doubt, the positive aspect of individuals actually playing mixed strategies is an open question of considerable interest.

Researchers' efforts to answer this question have naturally focused on settings where the use of mixed strategies is most compelling: the repeated play of games which have a unique mixed strategy Nash equilibrium. The value of "being unpredictable" is readily seen in examples such as serves in tennis, "bluffing" in poker, and whether or not a tax authority audits a tax payer. A common approach in this literature is to test whether the players' action choices are consistent with the mixed strategy equilibrium. Some studies using controlled experiments with human subjects have the advantage of knowing the payoff functions, and test whether choice frequencies agree with the equilibrium strategies and whether players' sequences of actions are serially independent (O'Neill, 1987; Binmore, Swierzbinski, and Proulx, 2001; Morgan and Sefton, 2002; Selten and Chmura, 2008). Other studies consider high-level sports competitions, such as soccer (Chiappori, Levitt, and Groseclose, 2002; Palacios-Huerta, 2003; Bareli, Azar, Ritov, Keidarlevin, and Schein, 2007) and tennis (Walker and Wooders, 2001), with the advantage of studying highly experienced players competing for high stakes and

---

[1]Along with generalizing the set of feasible actions to the set of mixed strategies, a player's payoff function is extended by setting its value to the expected payoff given a profile of mixed strategies, commonly referred to as the expected utility property.

the disadvantage of unknown payoff functions.[2] These studies focus on testing the serial independence of action choice and the equilibrium implication of equal payoffs across action choices. Some of the most prominent and recurring results for both types of studies are that aggregated action frequencies across players agree with the equilibrium mixed strategies but individual action frequencies do not, and for many individuals action choices are serially correlated violating the independence prediction.

To reconcile these issues of serial correlation and heterogeneity, several studies (Ochs, 1995; Bloomfield, 1994; Shachat, 2002; Noussair and Willinger, 2011) conduct laboratory experiments using the same type of games but directly elicit mixed strategies by obligating players to select a probability distribution over actions.[3] Elicited strategies in these experiments exhibit various distinct patterns. Some subjects choose pure strategies almost exclusively, some choose strictly mixed strategies almost exclusively, and others use both types of strategies – usually in long sequences. Also, certain mixed strategies are often quite focal, such as choosing equal probability weight on a subset of actions rather than the Nash equilibrium proportions. Naive interpretation of these results suggests a clear distinction between play that is purposely unpredictable and play that is a pure best response to changing forecasts of an opponent's action (Nyarko and Schotter, 2002). A more cautious interpretation is that subjects may eschew the randomizing device provided by the experimenter and instead internally randomize, or perhaps subjects choose strictly mixed strategies due to the experimenter effect of the novel elicitation method. Clearly a less invasive method to detect mixed strategy play would be valuable.

In this study we propose a hidden Markov model (HMM) to detect whether observed action choices are the result of pure or mixed strategies play in repeated two-person finite action games.[4] There are three key ideas in our formulation: (1) we treat the strategy a player follows as a latent state and the action played as the observable output from the latent strategy; (2) the set of possible latent states is a discrete subset of all possible mixed strategies containing pure strategies, Nash equilibrium or minimax strategies, and focal mixed strategies; and (3) a player switches the latent strategy he follows according to a first order Markov process. We then demonstrate the ability of the model by applying it to a new experimental data set we collect. In our experiment, each human subject repeatedly plays a 2×2 game against a computer player that follows its mixed strategy equilibrium. Some subjects play a zero-

---

[2]The action sets are typically comprised of simple actions, e.g., {serve left, serve right} and {defend left, defend right}. The payoffs are assumed to be the probability of winning the task and these probabilities will differ based upon both the comparative skills between players and the relative strengths a player has for each action.

[3]For example, Shachat (2002) adopts a game with four actions, each identified by a different color, for each player. Each player must fill a box with 100 cards in any combination of the four colored card types, and then one card is selected at random to determine the action played.

[4]See Rabiner (1989) for a classic introduction to hidden Markov models.

sum game and others an unprofitable game.[5] The estimated HMMs reveal several interesting results, including: (1) significant amounts of both pure and mixed strategy play; (2) the focal equiprobable mixed strategy is played more often than the Nash equilibrium strategy; (3) low transition probabilities between mixed and pure latent strategies; (4) dynamic adjustments in the types of strategies players follow over time; and (5) appreciable rates of both mixed and pure strategy play in the limiting distributions of the HMMs (interpreted as the long run equilibrium of play). We then extend the HMM from a statistical framework for evaluating hypotheses to one for forecasting action choice and assess its predictive accuracy.

2 A HMM of switching strategies

Consider an experiment in which we observe M pairs of subjects, each playing T periods of the same 2×2 normal form game. Often games like this are described by a two-by-two table, and for familiarity purposes we denote one subject's player role as Row and the other as Column. We label each player role's two possible actions Left (L) and Right (R), and express a subject's mixed strategy as the probability of playing L. Of particular interest is when the game has a single Nash equilibrium and it is in strictly mixed strategies, although our framework is not restricted to study only such cases. Three factors confounding the analysis data generated by this type of process are the latency of players' mixed strategies, the heterogeneity of strategy adoption across subjects, and variation of adopted latent strategies over the course of repeated play. In this section, we present a model that accommodates and allows estimation of these confounds.

Consider the following HMM for a fixed player role. The state space S is an n-element subset of the subject i's possible mixed strategies. Denote $s_{i,t} \in S$ for the strategy used by subject i in period t, S_i is the set of all possible T element sequences of mixed strategies for i with typical element s_i, and let $\mathbf{s}$ be the collection of s_i for all M subjects in a given player role. Let $y_{i,t}$ denote subject i's realized action in period t, y_i is the corresponding T element sequence of i's observable actions, and $\mathbf{y}$ is the collection of y_i for all M subjects. View $\{\mathbf{y}, \mathbf{s}\}$ as the output of the HMM.

The probability structure of the HMM has three elements. First, the n-element vector B for which the element B_j is the probability a subject chooses action Left, i.e. the mixed strategy, if he is in state j. We will provide two analyses which differ in how we specify B. In one approach we consider B as known *a priori*, and S and B are redundant notation. Usually, in this approach, B contains the two pure strategies, other strategies suggested by theory such

---

[5]An unprofitable game is one in which the minimax and Nash equilibrium solutions are distinct but yield the same expected payoff for each player.

as Nash equilibrium or minimax, and other focal strategies. In the second approach we treat the elements of B as unknown parameters – the state dependent mixed strategies. The second element of the structure, π, is the initial multinomial probability distribution over S. The third element, P, is the $n \times n$ transition probability matrix. The element P_{jk} is the probability a subject adopts strategy k in period t conditional upon having adopted strategy j in period $t - 1$.

The likelihood function of (B, π, P) is

$$L(B, \pi, P|\mathbf{y}, \mathbf{s}) = \Pr(\mathbf{y}, \mathbf{s}|B, \pi, P).$$

Rewriting this likelihood in terms of the marginal distributions of $\mathbf{y}$ and $\mathbf{s}$ gives us

$$L(B, \pi, P|\mathbf{y}, \mathbf{s}) = \Pr(\mathbf{y}|\mathbf{s}, B, \pi, P) \cdot \Pr(\mathbf{s}|B, \pi, P).$$

Next, we assume that the marginal distribution of $\mathbf{y}$ conditional on $\mathbf{s}$ is independent of π and P. In other words, once the state is realized then the probability of a Left action relies solely on the mixed strategy of the current state. Also, by the specification of the HMM, $\mathbf{s}$ is independent of the state dependent mixed strategies B. This allows us to restate the previous likelihood function as

$$L(B, \pi, P|\mathbf{y}, \mathbf{s}) = \Pr(\mathbf{y}|\mathbf{s}, B) \cdot \Pr(\mathbf{s}|\pi, P).$$

Since the sequence of states for each subject is unobservable, we evaluate the likelihood by integrating over the set of all possible sequences

$$L(B, \pi, P|\mathbf{y}, \mathbf{s}) =$$
$$\prod_{i=1}^{M} \sum_{\mathbf{s} \in \mathbf{S}} \pi(s_{i,1}) B_{s_{i,1}}^{I\langle y_{i,1}\rangle} (1 - B_{s_{i,1}})^{1 - I\langle y_{i,1}\rangle} \prod_{t=2}^{T} P_{s_{i,t-1}, s_{i,t}} B_{s_{i,t}}^{I\langle y_{i,t}\rangle} (1 - B_{s_{i,t}})^{1 - I\langle y_{i,t}\rangle},$$

where $I\langle y_{i,t}\rangle$ is an indicator function which equals one for the action Left and zero for the action Right. As T grows, the number of calculations needed to evaluate this likelihood quickly becomes computationally impractical. We describe the Bayesian approach we take to estimate the HMM, although one could proceed down a frequentist path of maximizing the expected likelihood function using some variation of the *EM* (expected maximum likelihood) algorithm.

In the Bayesian analysis, we first factor the joint posterior distribution of the unknown HMM parameters and unobserved states $\mathbf{s}$ into the product of marginal conditional posterior distributions. Then we evaluate these marginal conditional posteriors through an iterative sampling procedure called the Markov Chain Monte Carlo (MCMC) method. MCMC is a

simple but powerful procedure in which the empirical distributions of the sampled parameters converge to the true posterior distributions. After convergence, iterative sampling is continued to construct empirical density functions. These are used to make inferences regarding the parameters of the hidden Markov models.

Consider the posterior density function on the realized unobserved states and HMM parameters $h(\mathbf{s}, B, P, \pi|\mathbf{y})$. First, express this joint density as the product of the marginal density of HMM parameters conditional on the observed action choices and unobserved states with the marginal density of the states conditional upon action choices

$$h(\mathbf{s}, B, P, \pi|\mathbf{y}) = h(B, P, \pi|\mathbf{s}, \mathbf{y})h(\mathbf{s}|\mathbf{y}).$$

We have already assumed that the transition matrix P and initial probabilities over states π are independent of the action choices and state contingent mixed strategies B, which allows us to state

$$h(\mathbf{s}, B, P, \pi|\mathbf{y}) = h(B|\mathbf{s}, \mathbf{y})h(P, \pi|\mathbf{s}, \mathbf{y})h(\mathbf{s}|\mathbf{y}).$$

This product of three conditional posteriors permits a simple Markov Chain procedure of sequentially sampling from these distributions. We start with some initial arbitrary values for the HMM parameters, $(B^{(l)}, P^{(l)}, \pi^{(l)})$ where $l = 0$. We create $\mathbf{s}^{(0)}$ by simulation using $P^{(0)}$ and $\pi^{(0)}$ without conditioning on $\mathbf{y}$. From these initial parameter values and the observed action sequences $\mathbf{y}$, we use a Gibbs sampling algorithm to generate an initial sample of state sequences $\mathbf{s}^{(1)}$. Then we make a random draw $P^{(1)}$ from the posterior distribution of P conditional on $\mathbf{s}^{(1)}$ and $\mathbf{y}$, and proceed similarly to make a random draw of $\pi^{(1)}$. We complete the iteration by making a random draw $B^{(1)}$ from the posterior of B conditional on $\mathbf{s}^{(1)}$ and $\mathbf{y}$. The key to the MCMC method is that as $l \to \infty$, the joint and marginal distributions of $B^{(l)}$, $P^{(l)}$, and $\pi^{(l)}$ converge weakly to the joint and marginal posterior distributions of these parameters (Geman and Geman, 1987). We now describe the details of each step in an iteration of the MCMC procedure.

Step 1: Sampling the state sequences $\mathbf{s}^{(l)}$

We begin by describing a Gibbs sampling technique for generating draws from the distribution of $\mathbf{s}^{(l)}$ conditional upon $\mathbf{y}$ and $(B^{(l-1)}, P^{(l-1)}, \pi^{(l-1)})$. The elements of s_i can be drawn sequentially for each t conditioning on the observed action choice $y_{i,t}$, the realized state in other periods, π, and P. Let $s_{i,\neq t}$ be the vector obtained by removing $s_{i,t}$ from the sequence s_i. Given $s_{i,\neq t}$, we express the conditional posterior distribution of $s_{i,t}$ as

$$\Pr(s_{i,t}^{(l)}|y_{i,t}, B^{(l-1)}, P^{(l-1)}, s_{i,\neq t}^{(l)}, s_{i,\neq t}^{(l-1)}) \propto \Pr(y_{i,t}|s_{i,t}^{(l)}, B^{(l-1)}) \cdot \Pr(s_{i,t}^{(l)}|P^{(l-1)}, s_{i,\neq t}^{(l)}, s_{i,\neq t}^{(l-1)})$$

with

$$\Pr(s_{i,t}^{(l)}|P^{(l-1)}, s_{i,\neq t}^{(l)}, s_{i,\neq t}^{(l-1)}) = \Pr(s_{i,t} = k|P^{(l)}, s_{i,t-1}^{(l)}, s_{i,t+1}^{(l-1)}).$$

Consequently, the conditional posterior probability of $s_{i,t} = k$ and $t > 1$ is

$$\Pr(s_{i,t}^{(l)} = k|\cdot) = \frac{\Pr(y_{i,t}|s_{i,t} = k, B_k^{(l-1)}) \cdot \Pr(s_{i,t} = k|P^{(l-1)}, s_{i,t-1}^{(l)}, s_{i,t+1}^{(l-1)})}{\sum\limits_{j=1}^{n} \Pr(y_{i,t}|s_{i,t} = j, B_j^{(l-1)}) \cdot \Pr(s_{i,t} = j|P^{(l-1)}, s_{i,t-1}^{(l)}, s_{i,t+1}^{(l-1)})},$$

and for $t = 1$

$$\Pr(s_{i,1}^{(l)} = k|\cdot) = \frac{\Pr(y_{i,1}|s_{i,1} = k, B_k^{(l-1)}) \cdot \Pr(s_{i,1} = k|\pi^{(l-1)}, s_{i,2}^{(l-1)})}{\sum\limits_{j=1}^{n} \Pr(y_{i,1}|s_{i,1} = j, B_j^{(l-1)}) \cdot \Pr(s_{i,1} = j|\pi^{(l-1)}, s_{i,2}^{(l-1)})}.$$

The state $s_{i,t}^{(l)}$ is determined by making a random draw from the uniform distribution on the unit interval, and comparing this draw to the calculated conditional probability of $s_{i,t}^{(l)}$.

Step 2: Sampling the transition matrix $P^{(l)}$ and $\pi^{(l)}$

The posterior distributions of P_{jk} and π depend only upon $\mathbf{s}^{(l)}$ and the priors. We specify the prior of π as a Dirichlet distribution $h(\pi; \alpha_1, \ldots, \alpha_n)$ where $\alpha_j = 1$, for $1 \leq j \leq n$. Similarly, we specify the prior of the jth row of P as a Dirichlet distribution $h(p_{j1}, \ldots, p_{jn}|\eta_{j1}, \ldots, \eta_{jn})$. In an experiment, we record the data from the true start of the HMM process, so we assume that the joint posterior is simply the product of these two marginal posteriors. The respective posteriors of $\pi^{(l)}$ and $P^{(l)}$ are

$$h(\pi|\mathbf{s}) \propto \Pr(\mathbf{s}|\pi)h(\pi; \alpha_1, \ldots, \alpha_n),$$

and

$$h(P_{j1}, \ldots, P_{jn}|\mathbf{s}) \propto \Pr(\mathbf{s}|P_{j1}, \ldots, P_{jn})h(P_{j1}, \ldots, P_{jn}; \eta_{j1}, \ldots, \eta_{jn}).$$

If ν_{0j} is the number incidences of $s_{i,1}^{(l)} = j$ in $\mathbf{s}^{(l)}$, and ν_{jk} is the count of transitions from state j to k in $\mathbf{s}^{(l)}$, then the conditional probabilities in the two posterior calculations are multinomial distributions

$$h(\pi|\mathbf{s}) \propto \pi_1^{\nu_{01}} \ldots \pi_{n-1}^{\nu_{0n-1}} \cdot \left(1 - \sum_{k=1}^{n-1} \pi_k\right)^{\nu_{0n}} h(\pi; \alpha_1, \ldots, \alpha_n)$$

and

$$h(P_{j1}, \ldots, P_{jn}|\mathbf{s}) \propto P_{j1}^{\nu_{j1}} \ldots P_{jn-1}^{\nu_{jn-1}} \cdot \left(1 - \sum_{k=1}^{n-1} P_{jk} \right)^{\nu_{jn}} h(P_{j1}, \ldots, P_{jn}; \eta_1, \ldots, \eta_n).$$

Since the Dirichlet distribution is the conjugate prior for the multinomial distribution, these posterior distributions are also Dirichlet distributions for which each shape parameter is the sum of its prior value and the respective count

$$h(\pi|\mathbf{s}) = h(\pi; \alpha_1 + \nu_{01}, \ldots, \alpha_n + \nu_{0n})$$

and

$$h(P_{j1}, \ldots, P_{jn}|\mathbf{s}) = h(P_{j1}, \ldots, P_{jn}; \eta_1 + \nu_{j1}, \ldots, \eta_n + \nu_{jn}).$$

Hence, we select $\pi^{(l)}$ and $P^{(l)}$ be taking random draws from these distributions.

Step 3: Sampling the state dependent mixed strategies B

For our initial approach to modeling the state dependent mixed strategies, we assume B corresponds to a known subset of S. In our Bayesian analysis this is equivalent to assuming a point prior on these strategies, and therefore there is no updating. So in our Gibbs sampling procedure we skip this step, and proceed to next iteration of the Gibbs sampler. Of course this is a rather strong prior to assume, and we should evaluate whether it is appropriate. Accordingly, we conduct an auxiliary analysis in which we assume a uniform prior of the set of all mixed strategies.

In the auxiliary analysis we proceed as follows. The priors of state dependent mixed strategies $B_1, \ldots, B_n$ are assumed independent of each other and of the Markov process governing the states. Given these assumptions, we can think of each B_j as a Bernoulli probability, and each Left (Right) action as a success (failure) when occurring in state j. The likelihood function is calculated as a binomial trial. Since it is the conjugate prior of the binomial, we assume the prior is a Beta distribution, denoted $\beta(B_j; \zeta_j; \gamma_j)$. We want a uniform prior as well, and that corresponds to setting the shape parameters ζ_j and γ_j to one.

The posterior distribution is simply

$$h(B_j|\mathbf{y}, \mathbf{s}^{(l)}) = \beta(B_j; \zeta_j + \kappa_{L,j}, \gamma_j + \kappa_{R,j}),$$

where $\kappa_{L,j}$ and $\kappa_{R,j}$ are the number of times the actions Left and Right, respectively, are chosen when in state j according to $\mathbf{s}^{(l)}$. The state conditional mixed strategies $B_j^{(l)}$, $j = 1, \ldots, n$, are randomly drawn from these Beta posterior distributions, completing an iteration of the

Gibbs sampler.

The Gibbs sampler is run for a large number of iterations until the empirical distribution of all the parameters has converged (Geweke, 1991). Then the sampling procedure is allowed to continue to run for another number of iterations to build up an empirical distribution that corresponds to the posterior distribution of the HMM parameters. It is from this empirical distribution that we conduct statistical inferences.

3   The experiment

We apply our HMM framework to a new experimental data set that provides a likely setting for mixed strategies, and particulary Nash equilibrium strategies, to be adopted. Additionally, our procedures allow us to estimate for one player role without the need to also simultaneously model the opposing role, because each human subject repeatedly plays against a computer player that follows its mixed strategy equilibrium. Each subject is informed that his opponent is a computer but is given no information regarding the computer's strategy. We adopt two different games in our experimental design, with each subject playing only one of the two games. One game is zero-sum and the other game is unprofitable.

3.1   The games

Our first game is a zero-sum asymmetric matching pennies game introduced by Rosenthal, Shachat, and Walker (2003). The normal form representation of this game is presented on the left side of Figure 1. The game is called Pursue-Evade because the Row player "captures" points from the Column player when the actions of the two players match, and the Column player "evades" a loss when the players' actions differ. In the game each player can move either Left or Right, and the game has a unique Nash equilibrium in which each player chooses Left with probability two-thirds. In equilibrium, Row's expected payoff is two-thirds, and correspondingly Column's expected payoff is negative two-thirds.

Our second game is an unprofitable game introduced by Shachat and Swarthout (2004) called Gamble-Safe. Each player has a Gamble action (Left for each player) which yields a payoff of either two or zero, and a Safe action (Right for each player) which guarantees a payoff of one. The normal form representation of this game is presented on the right side of Figure 1. The Gamble-Safe game has a unique Nash equilibrium in which each player chooses the Left action with probability one-half, and each player earns an expected equilibrium payoff of one. Right is the minimax strategy for both players with a guaranteed payoff of one. Aumann (1985) argues that the Nash equilibrium prediction is not plausible in such an unprofitable game because its adoption assumes unnecessary risk to achieve the corresponding

Pursue-Evade Game				Gamble-Safe Game		

(The figure is a pair of payoff matrices.)

Pursue-Evade Game

Column Player

		Left	Right
Row Player	Left	1 , -1	0 , 0
	Right	0 , 0	2 , -2

Gamble-Safe Game

Column Player

		Left	Right
Row Player	Left	2 , 0	0 , 1
	Right	1 , 2	1 , 1

Figure 1: The experimental games

Nash equilibrium payoff. For example, imagine Row has Nash equilibrium beliefs and best responds by playing the Nash strategy. Row's expected payoff is one. However, suppose Column instead adopts his minimax strategy Right. This reduces Row's expected payoff to one-half. Row could avoid this risk by simply playing the minimax strategy. This aspect makes the Gamble-Safe game a more challenging test for the hypothesis of mixed strategy play than the zero-sum Pursue-Evade game.

3.2 Subject recruitment and experiment protocol

We conducted six experiment sessions in the Finance and Economics Experimental Laboratory (FEEL) at Xiamen University during December 2011. A total of 110 undergraduate and masters students participated in the experiment, with each session containing between 12 and 22 subjects. 54 subjects were assigned to the Persue-Evade game treatment, and 56 subjects were assigned to the Gamble-Safe game treatment. Subjects were evenly divided into Row and Column player roles within each treatment. FEEL uses the ORSEE online recruitment system for subject recruitment (Greiner, 2004), and at the time of the experiment approximately 1400 students were in the subject pool. A subset of students from the subject pool were invited to attend each specific session, and these students were informed that they would receive a 10 Yuan show-up payment and have the opportunity to earn more money during the experiment. Further, the invitation stated that the session would last no more than two hours.

Upon arrival at the laboratory, each subject was seated at a computer workstation such that no subject could observe another subject's screen. Subjects first read instructions detailing how to enter decisions and how earnings were determined.[6] Then, 200 repetitions of the game were played. For the Pursue-Evade game, Column subjects were initially endowed with a balance of 260 tokens each, and Row subjects none. Each token was worth one-third

---

[6]The instructions are available at `http://www.excen.gsu.edu/swarthout/HMM/`

of a Yuan. Each subject's total earnings consisted of the 10 Yuan show-up payment plus the monetary value of his token balance after the 200th repetition. While a mathematical possibility, no Column subjects in the Pursue-Evade game went bankrupt.

The experiment was conducted with a Java software application created at the Georgia State University Experimental Economics Center (ExCEN) that allows humans to play normal form games against computerized algorithms. At the beginning of each repetition, each subject saw a graphical representation of the game on the screen. Each Column subject's game display was transformed so that he appeared to be a Row player. Thus, each subject selected an action by clicking on a row, and then confirmed his choice. After the repetition was complete, each subject saw the outcome highlighted on the game display, as well as a text message stating both players' actions and his own earnings for that repetition. Finally, each subject's current token balance and a history of past play were displayed at all times. The history consisted of an ordered list with each row displaying the repetition number, the actions selected by both players, and the subject's payoff from the specific repetition.

3.3 Data summary

We begin the summary of the experimental data by providing views of the joint distribution of the proportion of Left play for each subject-computer pair, while conditioning on whether the data are from the first 100 or last 100 repetitions. Figures 2 and 3 present these views for the Pursue-Evade and Gamble-Safe treatments, respectively. In each of these figures, the x-axis is the proportion of Left play for the Column player and the y-axis is the proportion of Left play for the Row player. Each arrow in the figures represents the play of a single human-computer pair, with the arrow tail representing the joint frequency of Left play in the first 100 repetitions and the arrow head representing the joint frequency of Left play in the final 100 repetitions. These arrows show the adjustments subjects make from the first half to the second half of play. We see that many arrows suggest substantial change in the human player frequency, but the changes do not trend in any one direction or uniformly towards the Nash equilibrium. Human play also displays greater dispersion and displacement from the Nash equilibrium than the computer opponents, suggesting nonconformity with the Nash equilibrium predictions.

Table 1 presents the means and standard deviations of subjects' frequencies of Left play by treatment and role. Recall that we have 2700 observations for the each role in the Pursue-Evade treatment and 2800 observations for each role in the Gamble-Safe treatment. Although the Row player mean is close to the Nash equilibrium proportion in both game treatments, the Nash equilibrium proportion is rejected for all four cases at any reasonable level of significance. In each of the four cases, subjects' proportions of Left play display too much variance to have
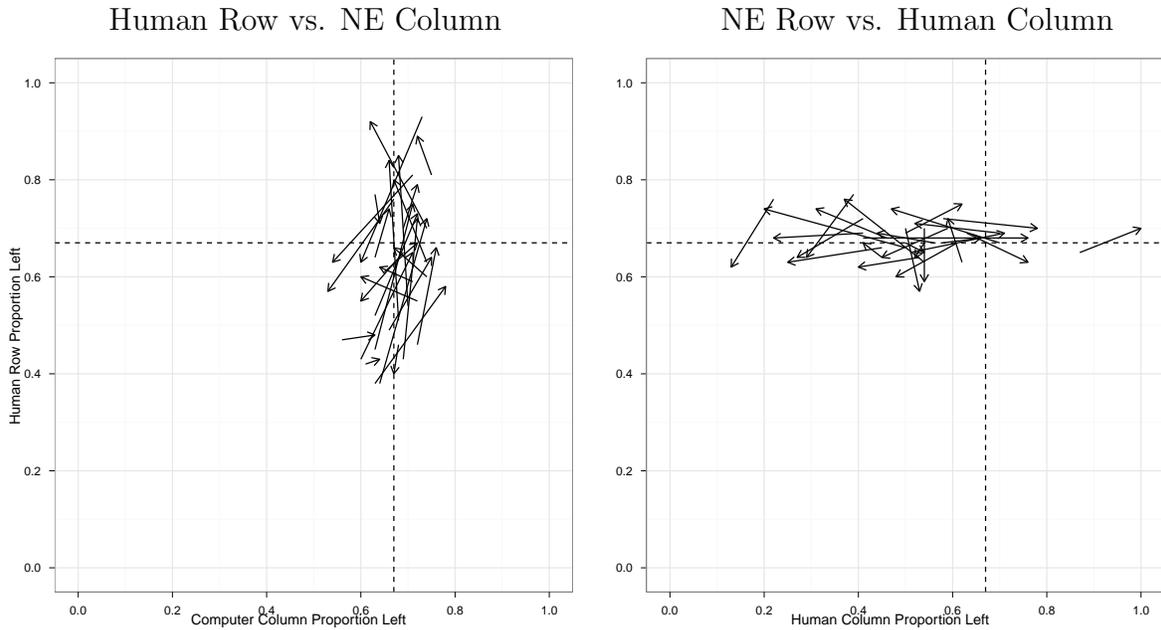
Figure 2: Pursue-Evade joint Left frequencies.
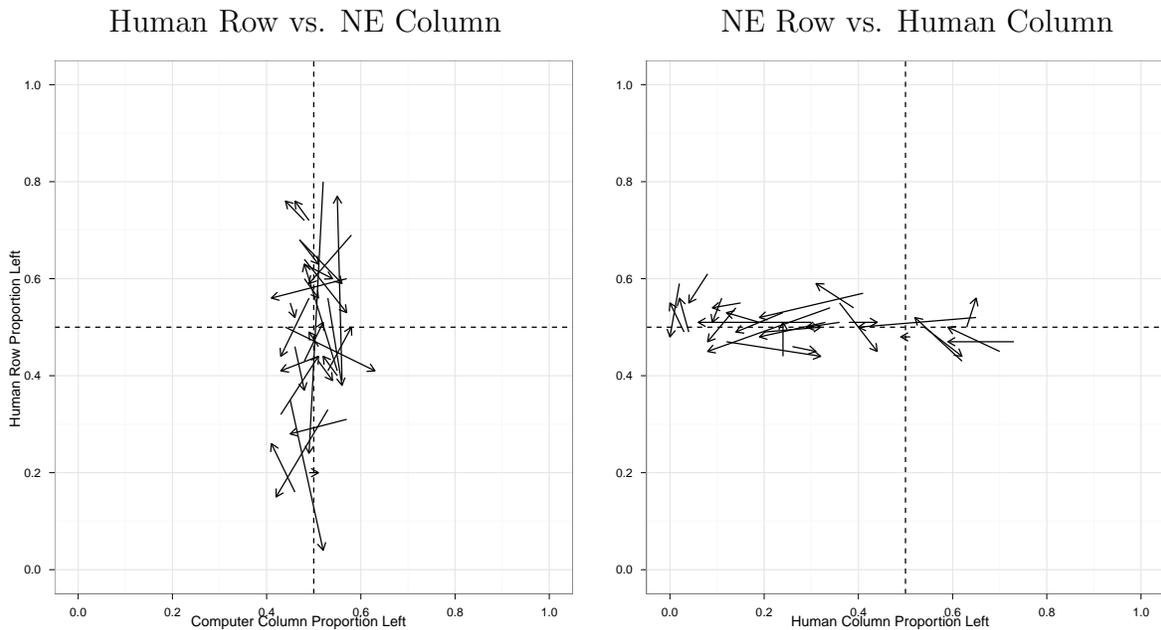
Figure 3: Gamble-Safe joint Left frequencies.

been generated by a common binomial process. In unreported χ^2 tests, we soundly reject this notion of homogeneity for all cases.

Rejecting the hypothesis of a single mixed strategy for each game and role, we present in Table 2 summary statistics for individual Pursue-Evade subjects and find evidence of heterogeneity and a lack of serial independence in choices. First, we present the number of times

Table 1: Aggregate Summary Statistics

Statistic	P-E Row	P-E Col	G-S Row	G-S Col
Average Left frequency	0.63	0.51	0.48	0.30
Standard deviation Left frequency	0.11	0.15	0.15	0.20
Nash equilibrium z-test statistic	-6.54	-25.06	-3.18	-30.20

the Left action is played in the first and second 100 repetitions. A two-tailed binomial test of the Nash equilibrium at the 95 percent level of confidence gives us critical regions of less than 58 and more than 76. We reject the Nash proportion of Left play for 13 (12) of the Row subjects during the initial (final) 100 repetitions, and we reject the Nash proportion for 21 (20) of the Column subjects during the initial (final) 100 repetitions.

Next, we evaluate whether the subjects' sequences of action choices are serially independent via a nonparametric runs test. The z-test statistic has a distribution approximate to the standard normal and is a function of the sequence length R, and the number of Left and Right sequences, r_L and r_R, respectively. Its value is

$$z = \frac{r_L + r_R - \frac{2 r_L r_R}{R} - \frac{1}{2}}{\sqrt{\left(\frac{2 r_L r_R (2 r_L r_R - R)}{R^2 (R-1)} \right)}}.$$

The null hypothesis of the test is that a subject's choices are independent realizations of a binomial random variable. We conduct a two-tailed test. Rejections from larger values of the test statistic indicate too many runs, and are symptomatic of negative serial correlation. Rejections from smaller values indicate too few runs, and are symptomatic of positive serial correlation. For the Row players, we reject serial independence for 10 subjects in the first half of the sample, and only 4 subject in the second half. For the Column players, the number of rejections is 14 and 10 for the first and second half, respectively. There is a notable bias with respect to the Column players; 22 out of 24 of the rejections come from z scores that are too negative and indicate strong positive serial correlation. This is consistent with the results found by Rosenthal et al. (2003) in the original study of the Pursue-Evade game, but atypical for other studies which often find negative serial correlation.

Table 3 presents a similar data summary for the individual subjects of the Gamble-Safe treatment. In this case, the Nash equilibrium mixed strategy is equiprobable, and the critical regions of the two-sided binomial tests are 39 or less and 60 or more Left action choices. For the Row players, the Nash hypothesis is rejected for 16 subjects in the first 100 repetitions and 15 in the second 100 repetitions. Correspondingly for the Column players, the Nash hypothesis is rejected for 25 subjects in the first half of repetitions and 21 players in the

Table 2: Pursue-Evade individual subject summary data.

	Row Player				Column Player			
	Rounds 1-100		Rounds 101-200		Rounds 1-100		Rounds 101-200	
Pair	Left Count	Runs Stat	Left Count	Runs Stat	Left Count	Runs Stat	Left Count	Runs Stat
1	$77^{n,e}$	-2.40^{i}	71^{e}	0.20	59^{n}	-4.44^{i}	45^{n}	-3.35^{i}
2	67^{e}	1.32	66^{e}	-0.20	50^{n}	6.23^{i}	53^{n}	3.47^{i}
3	$77^{n,e}$	-0.12	85^{e}	-0.60	51^{n}	1.01	71^{e}	0.69
4	49^{n}	-1.60	$65^{n,e}$	-1.22	$22^{n,e}$	-2.45^{i}	$13^{n,e}$	-2.08^{i}
5	59	1.17	63^{e}	-0.35	56^{n}	-1.28	44^{n}	-0.26
6	$39^{n,e}$	0.93	$37^{n,e}$	1.81	53^{n}	1.85	40^{n}	1.47
7	62^{e}	1.04	68^{e}	-0.58	41^{n}	-0.49	76^{e}	-2.35^{i}
8	47^{n}	0.44	48^{n}	-0.19	54^{n}	-3.58^{i}	52^{n}	0.22
9	51^{n}	-4.82^{i}	$84^{n,e}$	1.18	$63^{n,e}$	-1.64	76^{e}	-2.35^{i}
10	65^{e}	-4.53^{i}	55^{n}	-1.73	$39^{n,e}$	-3.08^{i}	$29^{n,e}$	-3.96^{i}
11	48^{n}	2.63^{i}	55^{n}	-1.12	46^{n}	-2.16^{i}	$31^{n,e}$	1.70
12	63^{e}	-2.29^{i}	75^{e}	0.94	54^{n}	-2.57^{i}	54^{n}	-3.78^{i}
13	60	-1.26	59	0.96	61^{e}	0.72	48^{n}	-1.39
14	64^{e}	-1.11	73^{e}	-0.36	54^{n}	-2.16^{i}	$37^{n,e}$	-1.64
15	$78^{n,e}$	0.79	66^{e}	-1.09	45^{n}	-3.76^{i}	41^{n}	-5.90^{i}
16	68^{e}	2.19^{i}	$88^{n,e}$	1.39	50^{n}	-2.01^{i}	54^{n}	-1.55
17	70^{e}	0.72	73^{e}	-1.64	50^{n}	-2.61^{i}	$20^{n,e}$	-3.16^{i}
18	46^{n}	-2.16^{i}	40^{n}	-2.51^{i}	70^{e}	-0.72	52^{n}	-1.39
19	51^{n}	2.02^{i}	65^{e}	1.00	41^{n}	-0.08	$22^{n,e}$	0.49
20	$80^{n,e}$	2.21^{i}	$78^{n,e}$	-0.09	45^{n}	-3.15^{i}	$25^{n,e}$	-6.86^{i}
21	51^{n}	0.61	64^{e}	-0.45	50^{n}	0.40	62^{e}	1.47
22	68^{e}	-0.12	$84^{n,e}$	-0.33	70^{e}	-1.68	47^{n}	0.64
23	64^{e}	-0.24	60	2.09^{i}	58	-1.39	$78^{n,e}$	-1.86
24	64^{e}	0.42	64^{e}	-0.24	58	1.71	66^{e}	-1.54
25	42^{n}	-0.97	43^{n}	-2.67^{i}	$87^{n,e}$	-2.98^{i}	$100^{n,e}$	—^{z}
26	76^{e}	0.42	57^{n}	-0.82	62^{e}	1.04	59	1.58
27	45^{n}	-6.19^{i}	75^{e}	-2.55^{i}	41^{n}	-2.57^{i}	$27^{n,e}$	-1.90

[n] Two-sided binomial test rejection of the NE proportion of 2/3 at the 5% level of significance.
[e] Two-sided binomial test rejection of equiprobable proportion at the 5% level of significance.
[i] Runs test rejection of serial independence at the 5% level of significance.
[z] Missing values due to inapplicability of test on data with zero variation.

second half of repetitions. Also, we see that 9 Column player subjects almost exclusively play the pure minimax strategy (over 90 times) in the last 100 repetitions, while there is only one such Row player. Further, we find evidence of serial correlation in many individuals' choice sequences. For the Row players, we reject serial independence for 12 and 9 subjects in the first and last 100 repetitions, respectively. For the Column players, serial independence is rejected for 12 subjects in the first half of repetitions and 5 subjects in the second half of repetitions.

Table 3: Gamble-Safe individual subject summary data.

Pair	Row Player				Column Player			
	Rounds 1-100		Rounds 101-200		Rounds 1-100		Rounds 101-200	
	Left Count	Runs Stat	Left Count	Runs Stat	Left Count	Runs Stat	Left Count	Runs Stat
1	43	0.61	39[m]	3.26	24[m]	−2.07[i]	14[m]	−0.88
2	43	−3.29[i]	51	1.21	19[m]	−0.91	12[m]	0.42
3	60[m]	−0.63	56	−1.08	36[m]	−3.29[i]	44	−5.36[i]
4	40	−1.05	44	−0.06	3[m]	−1.54	0[m]	—[z]
5	72[m]	0.17	76[m]	0.97	26[m]	0.66	31[m]	2.41[i]
6	50	−1.81	41[m]	−2.57[i]	70[m]	−4.80[i]	59	−0.49
7	41	0.13	63[m]	−1.43	15[m]	1.39	9[m]	−0.24
8	60[m]	−3.77[i]	56	−2.50[i]	63[m]	−1.86	65[m]	−0.77
9	16[m]	−2.22[i]	26[m]	−6.41[i]	33[m]	−2.78[i]	19[m]	−3.54[i]
10	39[m]	−0.12	77[m]	−3.25[i]	36[m]	−1.33	29[m]	−1.27
11	72[m]	3.42[i]	76[m]	−1.24	62[m]	−1.73	52	1.02
12	64[m]	3.48[i]	53	4.27[i]	24[m]	0.70	24[m]	−0.69
13	33[m]	−3.46[i]	15[m]	−2.59[i]	41	−1.33	19[m]	1.06
14	32[m]	−0.81	44	2.19[i]	51	−4.42[i]	49	−0.20
15	56	−1.89	44	−0.47	14[m]	1.23	8[m]	0.90
16	44	0.96	41	0.13	11[m]	0.22	9[m]	1.02
17	41	1.17	50	1.81	73[m]	−2.15[i]	59	−0.08
18	35[m]	−4.75[i]	4[m]	−6.50[i]	34[m]	−0.87	8[m]	0.90
19	31[m]	−0.89	28[m]	−1.58	54	2.09[i]	62[m]	0.83
20	56	−2.10[i]	38[m]	−0.67	4[m]	−2.33[i]	2[m]	0.24
21	68[m]	−0.35	59	0.96	8[m]	−3.30[i]	4[m]	0.44
22	46	−2.77[i]	37[m]	−1.64	2[m]	0.24	0[m]	—[z]
23	20[m]	−2.21[i]	20[m]	−0.32	20[m]	0.00	32[m]	−1.28
24	69[m]	0.05	59	−0.70	12[m]	0.42	32[m]	1.04
25	46	−0.74	49	1.01	65[m]	−2.10[i]	40	−2.51[i]
26	68[m]	2.42[i]	63[m]	0.30	39[m]	−0.12	31[m]	−1.60
27	80[m]	−4.74[i]	24[m]	−2.62[i]	38[m]	−2.37[i]	44	−1.49
28	63[m]	−0.13	60[m]	0.21	30[m]	−2.64[i]	6[m]	−3.03[i]

[m] Two-sided binomial test rejection of equiprobable proportion at the 5% level.
[i] Runs test rejection of serial independence at the 5% level of significance.
[z] Missing values due to inapplicability of test on data with zero variation.

4 Results of the HMM statistical analysis

In this section we present the estimated HMMs for the Pursue-Evade and Gamble-Safe treatments. First we report the means and variances of the posterior distributions of the transition probability matrices and the initial distributions over states. The estimates reflect adoption of both pure and mixed strategies and characterize the switching between latent strategies. We

then use these estimates to generate a description of the dynamics of the latent mixed strategy evolution. Finally, we provide an assessment of the robustness of some of our assumed priors.

4.1 Pursue-Evade game

For the Pursue-Evade game, we restrict the latent state space S to contain four elements. We treat the corresponding vector of state dependent mixed strategies B as fixed and known, and the four elements are the pure Right strategy (PR), the focal equiprobable mixed strategy (EM), the Nash equilibrium strategy (NE) of two-thirds, and the pure Left strategy (PL). Specifically, we assume a point prior of $B = (0, 0.5, 0.67, 1)$. Using this point prior we estimate the HMM using the MCMC method.

We run the the Gibbs sampler for 10,000 iterations. Using the last 5000 iterations, we establish that the empirical density functions have converged by applying the Geweke test (Geweke, 1991). Then we use these last 5000 iterations to make statistical inferences. Table 4 presents the estimated means and standard deviations of the transition probabilities between states, the same for the initial probabilities over state posteriors, and the calculated limiting distributions of the Markov chains for both Row and Column.

Table 4: Estimated transition matrices, initial and limiting distributions of Pursue-Evade game

	Row Player				Column Player			
	PR_{t+1}	EM_{t+1}	NE_{t+1}	PL_{t+1}	PR_{t+1}	EM_{t+1}	NE_{t+1}	PL_{t+1}
PR_t	0.75 (0.038)	0.145 (0.037)	0.051 (0.024)	0.054 (0.024)	0.752 (0.029)	0.204 (0.03)	0.025 (0.015)	0.018 (0.01)
EM_t	0.025 (0.009)	0.95 (0.012)	0.013 (0.006)	0.013 (0.006)	0.095 (0.028)	0.879 (0.033)	0.019 (0.012)	0.007 (0.005)
NE_t	0.005 (0.002)	0.007 (0.003)	0.939 (0.014)	0.05 (0.013)	0.012 (0.008)	0.021 (0.014)	0.96 (0.022)	0.007 (0.005)
PL_t	0.022 (0.011)	0.023 (0.011)	0.218 (0.035)	0.737 (0.034)	0.039 (0.017)	0.034 (0.016)	0.031 (0.015)	0.896 (0.026)
π	0.082 (0.063)	0.614 (0.13)	0.193 (0.12)	0.111 (0.08)	0.043 (0.040)	0.161 (0.127)	0.735 (0.141)	0.061 (0.057)
Limiting Distribution	0.050	0.274	0.548	0.128	0.178	0.385	0.356	0.080

Note: standard deviations are in parentheses.

Our estimation of the initial distribution over states is presented in the the fifth numeric row of Table 4. For both roles we find initial play has a high rate of mixed strategy play. Row players predominately follow the EM (61%), while the Column players predominantly

follow the NE (74%). Interestingly, this is quite different from the limiting distribution of the estimated transition matrices, which we can interpret as the long run steady state of the HMM. For the Row player, the mode of the limiting distribution is the NE (55%), while for the Column player both EM and NE are roughly equally likely, with probabilities of 39% and 36%, respectively. Clearly there is movement of strategy adoption over time.

Some aspects of these dynamics can be seen by inspection of the estimated transition probabilities, given in the first four numeric rows of Table 4. Large values on the main diagonals and corresponding small values on the off-diagonals indicate strong inertia in adopting new strategies. There are some interesting patterns when there is a transition between strategies. Consider the Row players first. When switching away from PR a player is almost three times as likely to switch to EM than either of the other two strategies. Likewise, when switching away from EM a player is twice as likely to switch to PR than either of the other strategies. There's a similar probabilistic cycle between NE and PL with much larger switching probabilities between them. The dynamic effects of these cycling tendencies can be seen in Figure 4, which presents time series of the estimated proportion of subjects using each of the four strategies.[7] The PL and the NE series tend to mirror one another, as do the PR and EM strategies – albeit with more noise.

The results for the Column players in the right hand side of Figure 4 are quite different. The use of NE steadily declines while the adoption of EM rises in the first 50 repetitions. Furthermore, we see a slow emergence of PR over the course of the experiment. The probabilistic cycle between the EM and PR strategies is evident by their sharp mirroring pattern.
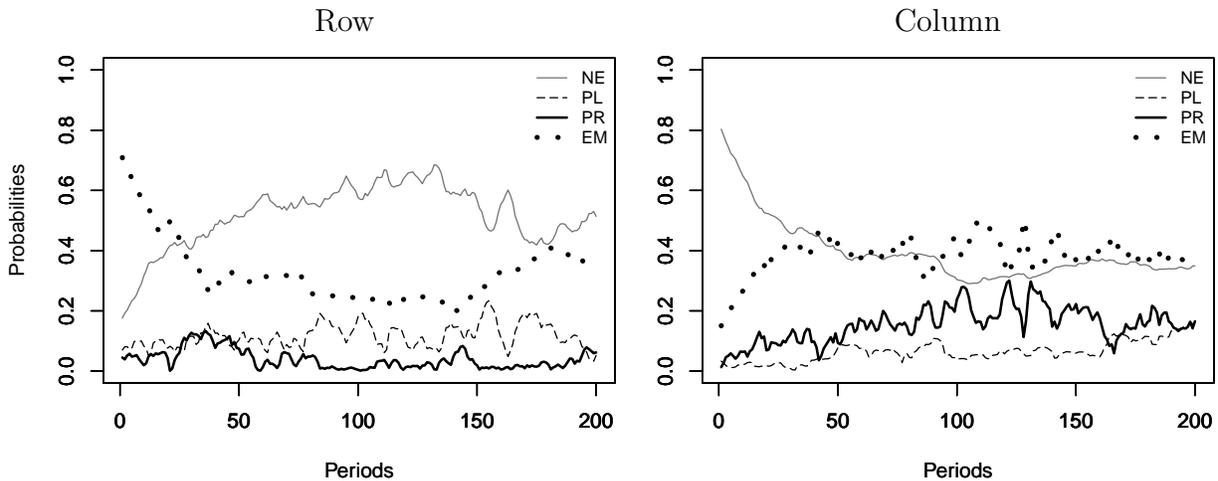

Figure 4: Strategy dynamics in Pursue-Evade game

Next we assess the appropriateness of our degenerate prior on B by conducting the MCMC

---

[7]For strategy j the estimated proportion of subjects using that strategy in a given round t is $\hat{j}_t = \frac{1}{27 \cdot 5000} \sum_{l=5001}^{10000} \sum_{i=1}^{27} I_{s_{i,t}^l = j}$.

estimation using a uniform Beta prior, $\beta(B_j; 1, 1)$, for each of the state dependent mixed strategies. We then sample from the posterior distributions to construct an empirical density function for each of the state dependent mixed strategies. In Figure 5 we present kernel smoothed plots of these approximations to posterior densities. Inspection reveals for the Row player the posteriors are sharply peaked and closely centered on our assumed four strategies, except for the NE and the posterior with a mode close to ³/₄ instead of ²/₃. For the Column player we see three out of four posteriors coincide with our assumed set. The one difference is the PR and the posterior with a mode of about 0.15.
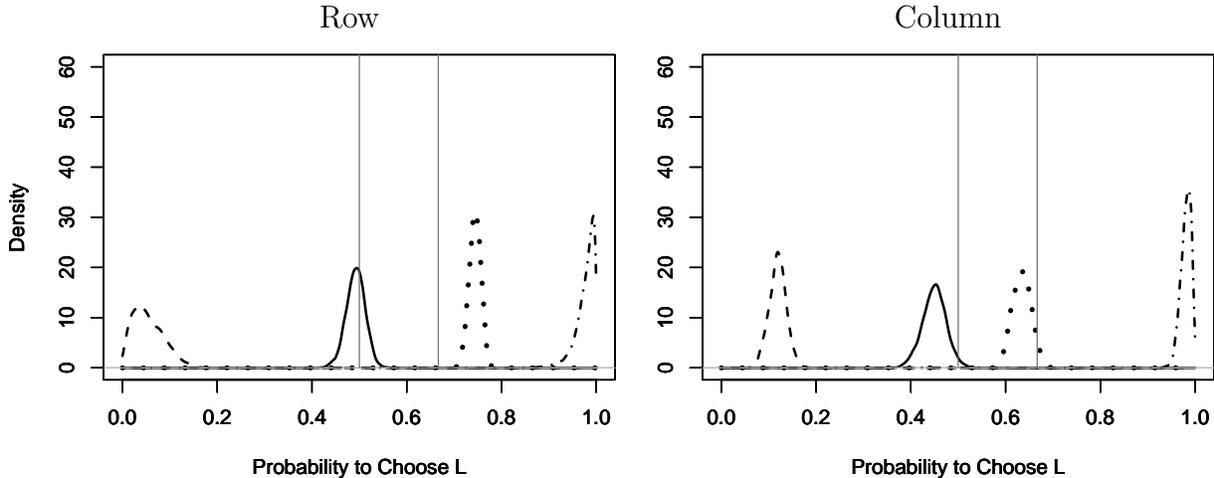
Figure 5: Posterior distribution of B in Pursue-Evade game

4.2 The Gamble-Safe game

We now turn our attention to the Gamble-Safe game. Here, we restrict the latent state space S to contain three elements. In our estimation we treat B as fixed and consisting of the elements PR (the minimax strategy), EM (which is also the NE strategy), and PL. We use the same parameters for the Gibbs Sampler as we used in analyzing the PE data.

For both the Row and Column player data sets we ran the Gibbs Sampler for 10,000 iterations, using the last 5000 iterations for inference after testing for convergence of the empirical densities with the Geweke test. The posterior means and standard deviations are reported in Table 5. Comparing the estimated initial distribution π to the limiting distribution suggests that an initial high probability of the mixed Nash strategy play reduces over time for both player roles. The change for the Column player is more dramatic as EM goes from 60% to 40%, and that reduction corresponds to a rise in the minimax strategy PR from 34% to 53%.

In contrast to the PE game, there is a segregation between mixed strategy and pure

strategy followers. Evidence of this is found in the estimated Markov transition matrices as we can see they almost fail to be irreducibile (roughly meaning we can always reach one state from another, even if it takes multiple transitions). The probability of continuing in the EM state is nearly one, indicating that once a subject follows the mixed strategy he is likely to do so for a large number of repetitions. Pure strategy adopters exhibit quite different patterns depending upon whether they are in the Row or Column role, in particular with respect to switching tendencies in the PL state. From the PL state, Row players transition to PR with 26% probability, while this transition probability is 79% for Column players.

Table 5: Estimated transition matrices, initial and limiting distributions of Gamble-Safe game

	Row Player			Column Player		
	PR_{t+1}	EM_{t+1}	PL_{t+1}	PR_{t+1}	EM_{t+1}	PL_{t+1}
PR_t	0.815 (0.027)	0.010 (0.020)	0.175 (0.016)	0.891 (0.010)	0.006 (0.010)	0.103 (0.007)
EM_t	0.003 (0.011)	0.988 (0.021)	0.009 (0.011)	0.007 (0.013)	0.985 (0.019)	0.008 (0.008)
PL_t	0.260 (0.031)	0.043 (0.031)	0.697 (0.037)	0.791 (0.031)	0.042 (0.031)	0.167 (0.044)
π	0.169 (0.084)	0.779 (0.094)	0.052 (0.046)	0.337 (0.098)	0.596 (0.103)	0.067 (0.052)
Limiting Distribution	0.203	0.660	0.137	0.527	0.404	0.059

Note: standard deviations are in parentheses.

Figure 6 presents the time series of the estimated proportion of subjects using each of the three latent strategies. Here we see the impact of the Markov transition probabilities that lead to inertia of the mixed strategy state and also the strong cycling tendencies of players between the Left and Right pure strategies. In the Row player figure, we see the EM strategy proportion has a smooth path that drops quickly from its initial level to its limiting value within the first 50 repetitions, after which it remains relatively constant. We also see the ragged mirroring pattern, indicating switching between the PR and PL strategies. We see similar features in the Column figure except that the EM shows a more gradual decline, and PR shows a corresponding gradual increase. This leads to the separation of the PR and PL strategies and allows us to see the clear short run switching between these strategies characterized by the jagged mirror relationship between their respective series.

We test the robustness of our point prior $B = [0, 0.5, 1]$, by estimating an HMM for which these state conditional strategies each have a uniform Beta prior. The kernel smoothed empirical density functions of the posteriors are presented in Figure 7 for both Row and
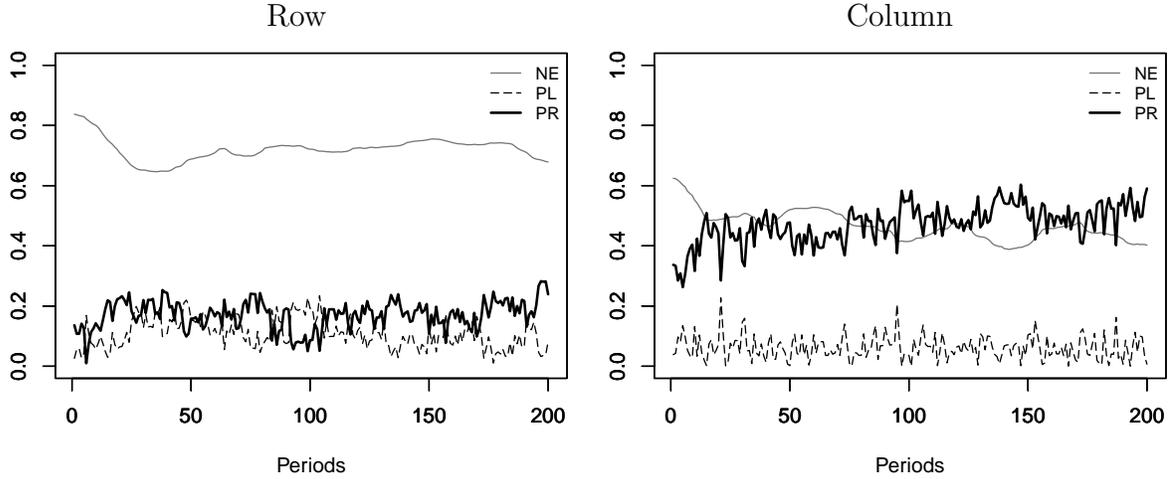
Figure 6: Strategy dynamics in Gamble-Safe game

Column players. For the Row player, the lower and middle posteriors are closer together than our assumed sets. For the Column player, the posteriors of the lower two state dependent strategies are shifted to the right of our assumed ones. We conjecture these shifts could come from erroneously assumed homogeneity of the strictly mixed strategy used by subjects. An alternative would be to increase the number of elements in S or to model the individuals' strictly mixed strategies coming from a hierarchical process.
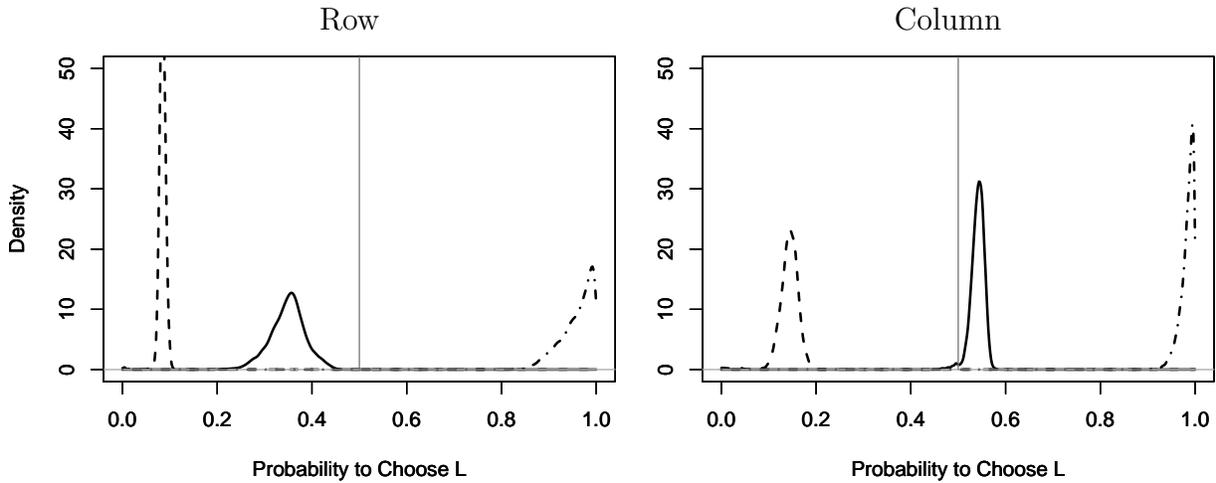

Figure 7: Posterior distribution of B in Gamble-Safe game

4.3 Forecasting realized actions

Until now our primary concern has been the estimation of when subjects adopt pure and mixed strategies, and our HMM's function has been to provide a statistical framework to test

20

theories about latent strategy choice. Now we explore the potential of the HMM to predict actions taken; a valuable capability in widespread applications from strategic maneuvers in military engagements, to knowing when a poker player is bluffing.

We first consider how well the estimated HMMs coincide with the observed proportions of Left play in our experimental data set. For this forecasting exercise of the experimental panel data set we calculate, for each game and role, the predicted proportion of Left play by the M subjects in period t, $\widehat{Left}_t$, by

$$\widehat{Left}_t = \frac{1}{M \cdot L} \sum_{l=1}^{L} \sum_{d=1}^{M} \sum_{j=1}^{N} (I_{s^l_{d,t}=j}) \beta_j.$$

Here L is the length of sequence of the Gibbs sampler we use for statistical inference. For our data sets this sequence is iterations 5001 to 10000. Figure 8 presents plots of the time series of the predicted and actual proportions of Left play. In all four settings the predictions track the trends in the actual data. Admittedly this is an in-sample forecasting exercise, but nonetheless still impressive, as minimizing forecast error is not the objective of our statistical inference exercise.

Out-of-sample forecasting is of more practical use and we can use the HMM for this purpose as well. We estimate, with 10,000 iterations of the Gibbs sampler, the HMM for both point and uniform Beta priors on B for the first 100 repetitions and use these estimates to make one-step-ahead forecasts of the last 100 repetitions. Let $\Psi = (P, \pi, B, s_{i,t})_{j=5001}^{10000}$ denote the realized draws of the Gibbs sampler for the last 5000 iterations of the MCMC algorithm that are used for statistical inference for the uniform Beta prior HMM. The predictive density of $s_{i,t}$ is obtained by simulation from the joint posterior sample Ψ as follows:

$$\hat{s}_{i,t}^{[j]} \sim p(s_{i,t}|P^{[j]}, \tilde{s}_{i,t-1}^{[j]}), j = 5001, \ldots, 10000. \tag{1}$$

We can use these sampled states for subject i to generate the following 5000 draws from the following marginal posterior sample

$$\hat{y}_{i,t}^{[j]} \sim p(y_{i,t}|\hat{s}_{i,t}^{[j]}, B^{[j]}), j = 5001, \ldots, 10000. \tag{2}$$

The average of the 5000 draws made according to Equation 2, denoted $\hat{y}_{i,t}$, is the prediction of $y_{i,t}$. Next we use $y_{i,t}$ to generate the posterior density $\tilde{s}_{i,t}$ by Bayes' Rule. This is substituted into Equation 1 to start the process of generating the prediction of $y_{i,t+1}$. To assess the accuracy of our forecast of the holdout sample, we calculate and report the Log-likehood
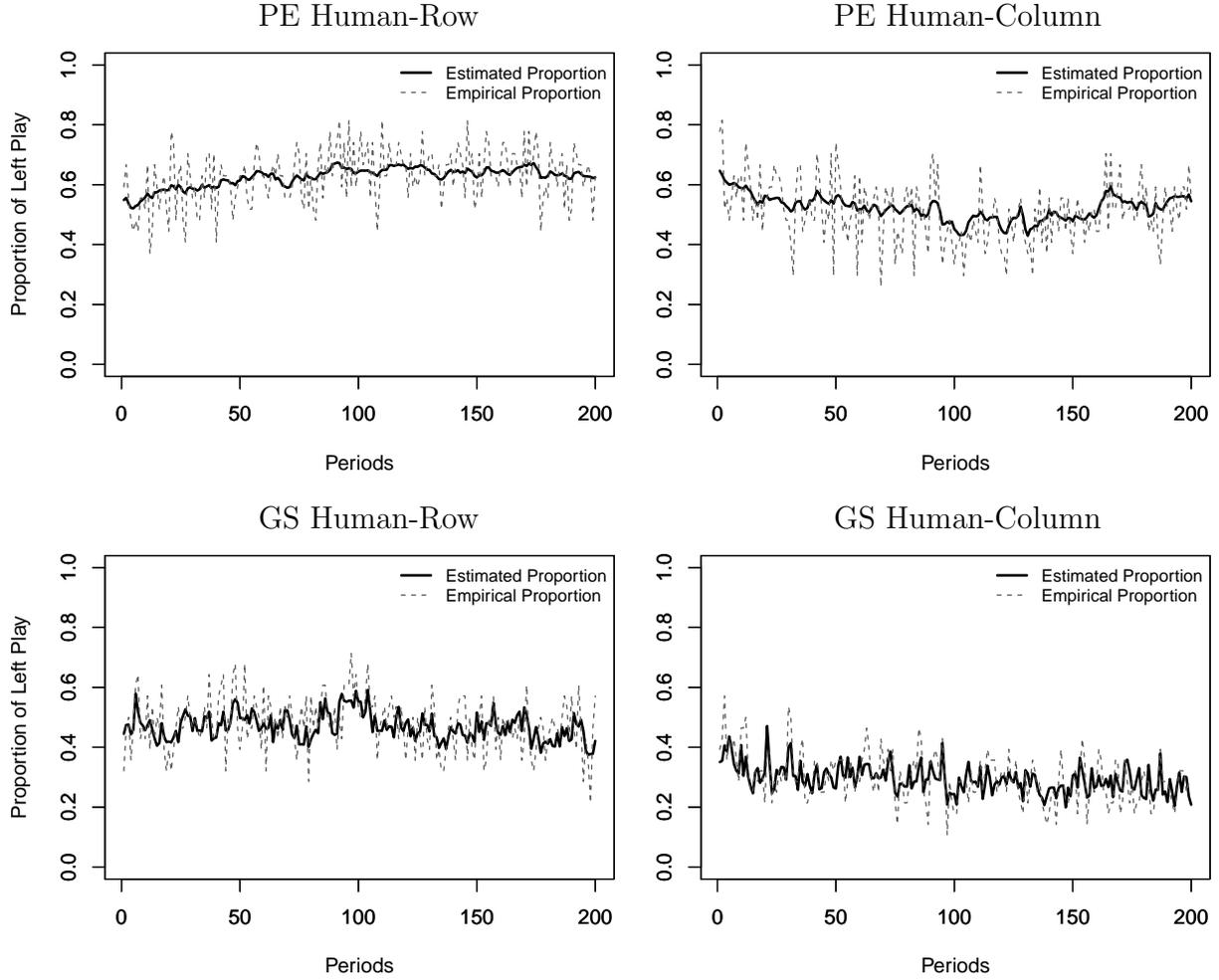
Figure 8: Actual and forecasted proportions of Left play

statistic

$$LL(y|\Psi) = \sum_{t=101}^{200} \sum_{i=1}^{M} \ln[I\langle y_{i,t}\rangle p(\hat{y}_{i,t}) + (1 - I\langle y_{i,t}\rangle)(1 - p(\hat{y}_{i,t}))].$$

We also report the Akaike information criterion statistic, which is $AIC(y, \hat{y}) = -2 \cdot (LL-$ number of model parameters).

In order to evaluate the ability of alternative models to predict the future actions in games, we compare the performances of one-step-ahead forecasting of the HMMs of point and uniform Beta priors (UHMM) on B against the alternatives of the Nash equilibrium strategy and individual-specific mixed strategies (IM) which are estimated by each subject's proportion of Left play in the first 100 repetitions.

We summarize the out-of-sample forecasting performance for each of the four models in Table 6. First, for the Row players in both game treatments the two HMMs outperform the two other models when we do and do not penalize for the number of parameters. For the

hold out sample of the Column players and not penalizing for the number of parameters, the IM model performs comparable to the two HMM models in the Pursue-Evade game, and the IM model performs comparable to the UHMM model (both of which outperform the HMM) in the Gamble-Safe game. However, when we penalize for increasing numbers of parameters we see the UHMM clearly outperforms the IM model. This suggests that our homogeneous dynamic model performs well on forecasting a population of game players, but also suggests that allowing for more individual heterogeneity could lead to even better out of sample forecasting performance.

Table 6: Out of sample forecasting performance

Treatment	Statistic	Row Player				Column Player			
		NE	IM	HMM	UHMM	NE	IM	HMM	UHMM
P-E Game	Loglik	−1748	−1736	−1723	−1710	−2058	−1775	−1776	−1771
	AIC	3497	3526	3475	3458	4117	3605	3583	3580
G-S Game	Loglik	−1941	−1916	−1887	−1865	−1941	−1429	−1517	−1436
	AIC	3884	3887	3789	3751	3885	2915	3050	2894

5   Discussion

We have introduced a HMM for the detection of pure and mixed strategy play in repeated games. We then applied this model to data from a new experiment in which human subjects repeatedly play against computer opponents that were programmed to play their part of the mixed strategy Nash equilibrium. We find that subjects do play both pure and mixed strategies, and switch between these over the course of play. Further, we find there is non-stationarity in the distribution of latent strategies over time. We observe a large movement from the initial distribution over strategies to those of the limiting distribution of the HMM. However, while the limiting distribution assigns probability to the subjects' NE strategy, the assigned probability is less than 1. Thus, for our data, we show that a mixed strategy Nash equilibrium is only partially self-enforcing. This is a new result in behavioral game theory, as previous studies have only considered the composite hypothesis that mixed strategy equilibria are both self-enforcing and also the limit point of the subjects' learning process.

Our primary interest has been modeling a population of players interacting in a game with known payoffs, however there are several natural extensions to our approach. First, we could focus on the modeling and forecasting of a single subject from the population. To do this, we likely need to allow more individual heterogeneity in the HMM. A first step would be to allow each player to have a set of individual-specific strict mixed strategies to follow. This

could done by allowing individual state dependent mixed strategies B_{is}, or by modeling these B_{is} as coming from a hierarchal structure characterized by a small set of hyperparameters. A second extension is to model strategic situations in the field, in which the game payoffs are not known because of unobserved individual heterogeneity. For example, soccer players making penalty kicks may vary in their strength of kicking left or right, and similarly goalies also may have unobservable differences in defending kicks to the left and right. In such cases, the HMM can help identify such payoffs and also describe the players' learning process regarding these latent payoff types.

The HMM as presented in this paper is currently more of a statistical description than a behavioral model derived from optimizing behavior. To become such a behavioral model, the transition probabilities must become an endogenous function of a player's expected payoffs for the differing latent strategy choices. One possible approach is to allow a player to form beliefs about an opponent's action and then best respond. The issue here is that in an expected utility world a mixed strategy is never a strict best response. However, if one takes the approach that uncertainty about an opponent's action is ambiguous – i.e., a player doesn't have the ability to form a unique prior – then an ambiguity-averse player may strictly prefer a mixed strategy over pure strategies.

References

Aumann, Robert J. (1985), "On the non-transferable utility value: A comment on the Roth-Shafer examples." *Econometrica*, 53, 667–678.

Bareli, M, O Azar, I Ritov, Y Keidarlevin, and G Schein (2007), "Action bias among elite soccer goalkeepers: The case of penalty kicks." *Journal of Economic Psychology*, 28, 606–621.

Binmore, Ken, Joe Swierzbinski, and Chris Proulx (2001), "Does minimax work? an experimental study." *Economic Journal*, 111, 445–64.

Bloomfield, Robert (1994), "Learning a mixed strategy equilibrium in the laboratory." *Journal of Economic Behavior & Organization*, 25, 411–436.

Chiappori, Pierre, Steven Levitt, and T. Groseclose (2002), "Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer." *American Economic Review*, 92, 1138–1151.

Geman, Stuart and Donald Geman (1987), "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." In *Readings in computer vision: issues, problems,*

principles, and paradigms (Martin A. Fischler and Oscar Firschein, eds.), 564–584, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Geweke, John (1991), "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments." Staff Report 148, Federal Reserve Bank of Minneapolis.

Greiner, Ben (2004), "An online recruitment system for economic experiments." In *Forschung und wissenschaftliches Rechnen* (Kurt Kremer and Volker Macho, eds.), volume 63 of *Ges. fur Wiss. Datenverarbeitung*, 79–93, GWDG Bericht.

Morgan, John and Martin Sefton (2002), "An experimental investigation of unprofitable games." *Games and Economic Behavior*, 40, 123–146.

Nash, John (1951), "Non-Cooperative Games." *The Annals of Mathematics*, 54, 286–295.

Noussair, Charles and Marc Willinger (2011), "Mixed strategies in an unprofitable game: an experiment." Working papers, LAMETA, Universtiy of Montpellier.

Nyarko, Yaw and Andrew Schotter (2002), "An experimental study of belief learning using elicited beliefs." *Econometrica*, 70, 971–1005.

Ochs, Jack (1995), "Games with unique, mixed strategy equilibria: An experimental study." *Games and Economic Behavior*, 10, 202–217.

O'Neill, Barry (1987), "Nonmetric test of the minimax theory of two-person zero-sum games." *Proceedings of the National Academy of Sciences, U.S.A.*, 84, 2106–2109.

Palacios-Huerta, Ignacio (2003), "Professionals play minimax." *Review of Economic Studies*, 70, 395–415.

Rabiner, Lawrence R. (1989), "A tutorial on hidden markov models and selected applications in speech recognition." *Proceedings of the IEEE*, 77, 257–286.

Rosenthal, Robert W., Jason Shachat, and Mark Walker (2003), "Hide and seek in arizona." *International Journal of Game Theory*, 32, 273–293.

Selten, Reinhard and Thorsten Chmura (2008), "Stationary concepts for experimental 2x2-games." *American Economic Review*, 98, 938–966.

Shachat, Jason (2002), "Mixed strategy play and the minimax hypothesis." *Journal of Economic Theory*, 104, 189–226.

Shachat, Jason and J. Todd Swarthout (2004), "Do we detect and exploit mixed strategy play by opponents?" *Mathematical Methods of Operations Research*, 59, 359–373.

Von Neumann, John (1928), "Zur theorie der gesellschaftsspiele." *Mathematische Annalen*, 100, 295–320.

Von Neumann, John and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton University Press.

Walker, Mark and John Wooders (2001), "Minimax play at Wimbledon." *American Economic Review*, 91, 1521–1538.