2-21-2011

Man versus Nash An experiment on the self-enforcing nature of mixed strategy equilibrium

Jason Shachat
Xiamen University

Todd Swarthout
Georgia State University

Lijia Wei
Xiamen University

Recommended Citation

Shachat, Jason; Swarthout, Todd; and Wei, Lijia, "Man versus Nash An experiment on the self-enforcing nature of mixed strategy equilibrium" (2011). *ExCEN Working Papers*. 69.
https://scholarworks.gsu.edu/excen_workingpapers/69

Man versus Nash

An experiment on the self-enforcing nature of mixed strategy equilibrium

Jason Shachat[*]
J. Todd Swarthout[†]
Lijia Wei[‡]

February 21, 2011

Abstract

We examine experimentally how humans behave when they play against a computer which implements its part of a mixed strategy Nash equilibrium. We consider two games, one zero-sum and another unprofitable with a pure minimax strategy. A minority of subjects' play was consistent with their Nash equilibrium strategy, while a larger percentage of subjects' play was more consistent with different models of play: equiprobable play for the zero-sum game, and the minimax strategy in the unprofitable game. We estimate the heterogeneity and dynamics of the subjects' latent mixed strategy sequences via a hidden Markov model. This provides clear results on the identification of the use of pure and mixed strategies and the limiting distribution over strategies. The mixed strategy Nash equilibrium is not self-enforcing except when it coincides with the equal probability mixed strategy, and there is surprising amounts of pure strategy play and clear cycling between the pure strategy states.

Keywords Mixed Strategy · Nash Equilibrium · Experiment · Hidden Markov Model

JEL Classification C92 · C72 · C10

1   Introduction

A Nash equilibrium is self enforcing because no player can strictly gain through unilateral deviation. Pure strategy Nash equilibria are generically strict in that a player's Nash equilibrium strategy is the unique best response when all other players follow their respective part of the Nash strategy profile; so if the player unilaterally deviates his payoff will be strictly lower. In contrast, mixed strategy Nash equilibria are generically weak (Harsanyi, 1973). Not only is a player's Nash equilibrium strategy a best response to all other players following their respective Nash equilibrium strategies, but so is any other mixture over the support of the player's equilibrium strategy. The non-uniqueness of best responses raises the question of whether a mixed strategy Nash equilibrium is self enforcing in actual decision making situations.

Repeated strictly competitive games, such as constant- and zero-sum games, provide an ideal environment in which to examine the behavioral robustness of a mixed strategy Nash equilibrium. We expect this

---

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

[†]Department of Economics, Georgia State University, *swarthout@gsu.edu*

[‡]Wang Yanan Institute for Studies in Economics, Xiamen University, *ljwei.wise@gmail.com*

class of games to provide the theoretical predictions the best chance of succeeding, as the Nash equilibrium and minimax solution concepts coincide, and randomizing appropriately over time insures against exploitation. Not surprisingly, there is a an extensive literature of such experimental studies, from which there are conflicting results. On the negative side, several studies reject hypotheses derived from mixed strategy Nash equilibrium theory (Brown and Rosenthal, 1990; Rosenthal, Shachat, and Walker, 2003; Shachat, 2002; Wooders, 2010; Levitt, List, and Reiley, 2010), and others find repeated play typically fails to converge to the Nash equilibrium (Mookherjee and Sopher, 1994; Erev and Roth, 1998; Selten and Chmura, 2008). However on the more positive side, others argue that while one can statistically reject the precise predictions of a mixed strategy Nash equilibrium, these predictions do qualitatively describe how subjects play (O'Neill, 1987; Binmore, Swierzbinski, and Proulx, 2001; Palacios-Huerta and Volij, 2008). Further, others find support for the hypothesis that learning – or more specifically, the adjustment of action choice proportions over time – converges to the Nash profile (Bloomfield, 1994; Mukherji, McCabe, and Runkle, 2000).

In a typical study of this literature, fixed pairs of human subjects repeatedly play a simple normal form game with a unique Nash equilibrium that is in mixed strategies. When there is failure to achieve equilibrium in such an experiment, one can't discern whether the adaptive joint dynamics of play fail to converge to the equilibrium (an instability property of the equilibrium), or if the equilibrium simply is not self enforcing. To independently test the self enforcing characteristic of an equilibrium, one could allow a single player to choose his action while ensuring that all other players follow their respective strategies of the Nash equilibrium strategy profile.

To isolate the self enforcing nature of a mixed strategy equilibrium from the issue of stability, we utilize a methodology that is a hybrid of experimentation with human subjects and computerized agent-based simulations. In our study a human subject repeatedly plays a 2×2 game against a computer player that follows its mixed strategy equilibrium. We adopt two different games in our experimental design, but each subject plays only one. One game is zero-sum, selected to give the theory the best chance of succeeding. The other game is a unprofitable one. An unprofitable game is one in which the minimax and Nash equilibrium solutions are distinct but yield the same expected payoff for each player. This is a more challenging environment for the Nash Equilibrium versus a zero-sum one, as discussed in Aumann and Maschler (1972) and Aumann (1985). We believe this design provides an ideal setting to investigate whether the mixed strategy Nash equilibrium is self enforcing, and to separate the extent it is self enforcing due to equilibrium beliefs (i.e., Nash equilibrium) or security concerns (i.e., minimax).

In the first part of the data analysis, we evaluate three alternative hypotheses regarding subject behavior. The first hypothesis is that subjects adopt their part of the mixed strategy Nash equilibrium. For most subjects we reject this hypothesis. The second hypothesis is that subjects adopt an equiprobable mixed

strategy. There is support for this hypothesis in the aggregated data across subjects in the zero-sum game; however, at the individual level there is too much variation in action choice proportions to conclude that subjects all play according to the equiprobable mixed strategy. The third hypothesis is that subjects who play the unprofitable game will play a pure strategy minimax strategy rather than their part of the mixed strategy Nash equilibrium. We find some support for this hypothesis. The conclusions of this part of the data analysis are that there is strong heterogeneity in subject play, and possible adjustments of play over time.

The second part of the data analysis estimates this strategic heterogeneity and dynamics with a hidden Markov model(HMM). The latent states of the HMM are a subset of the set of all mixed strategies: the Nash strategy, the equiprobable strategy, and the two pure strategies. This model also consists a Markov transition matrix which governs how subjects switch between latent strategy states. The estimated HMM provides several interesting results for the zero-sum game. First, we find that strategy switching is quite slow and in the limiting distribution of the Markov transition matrix the equiprobable strategy is adopted over fifty percent of the time, the Nash equilibrium strategy less than seventeen percent, and a surprising over thirty percent of the time players use pure strategies. There is also an asymmetry across the player roles, with one player adopting one of the pure strategies over twenty-two percent in the limiting distribution. In the unprofitable game, the Nash equilibrium strategy is the equiprobable strategy. Here we find for one of the player roles, the Nash strategy has a strong attraction. It's estimated that less than sixty percent of the players initially adopt the Nash strategy, which rises to over eighty percent by the end of the experiment, and is over ninety percent in the limiting distribution. There is a role effect for this game as well. The other player splits its play primarily between the Nash and minimax strategies. Furthermore, in both games we find a strong tendency for players adopting pure strategies to transition to the other pure strategy rather than a mixed strategy, generating cyclic play. We believe this is the first time such dynamic models of discrete heterogeneity have been applied to this type of data.

Our current study contributes to a second literature, experiments with human-computer interaction in repeatedly played games with a unique mixed strategy equilibrium. Such hybrid studies of human-agent interaction are a natural extension suggested by influential survey articles such as Duffy (2006) and Richiardi, Leombruni, and Contini (2006).[1] Several studies have used the approach of a human subject playing zero-sum games against a computer implementing a mixed Nash equilibrium strategy (Lieberman, 1962; Messick, 1967; Fox, 1972). These studies have found that the human play does not conform to the Nash equilibrium strategies, but these studies substantially differ from ours as they informed subjects they were playing against

---

[1]Furthermore, human interaction with computerized agents is an emerging important phenomenon. For example, in 2009 the majority of all trades in US asset markets involve algorithm traders (Sussman, Tabb, and Lati, 2009).

a computer (but did not inform about the computer's strategy), and sometimes didn't fully reveal the payoff matrix. Other experiments have shown that humans are generally are able to exploit computers in similar games, when the computers follow non-equilibrium stationary mixed strategies (Lieberman, 1962; Fox, 1972; Shachat and Swarthout, 2004). Another part of this literature reports experiments with humans playing against computer agents which follow various adaptive algorithms (Shachat and Swarthout, 2008; Messick, 1967; Coricelli, 2001; Duersch, Kolb, Oechssler, and Schipper, 2010; Spiliopoulos, 2008).

Finally, our study sheds additional light on previous research showing that in non-zero-sum games which have a unique Nash equilibrium in mixed strategies, subjects' play depends significantly on the magnitude of their own potential game payoffs (Ochs, 1995; McKelvey, Palfrey, and Weber, 2000; Willinger and Noussair, 2003). In this present study, we consider a different own-payoff-function effect by comparing the behavior in a zero-sum game to that in an unprofitable game. Morgan and Sefton (2002) and Shachat and Swarthout (2004) are the only other studies we know of which conduct experiments with unprofitable games. Morgan and Sefton (2002) find that in 3×3 unprofitable games, neither solution concept accurately describes play in their experiments.

Next we describe our experimental design. Then in section four we present a data analysis assessing the different alternative hypotheses of play. In the penultimate section, we present our hidden Markov model of heterogeneity and strategy switching. Finally, we provide some concluding remarks.

2   Experimental design

There are three elements of the experimental design: the choice of games, the subject pool composition and recruitment practice, and then the procedures used to carry out the experiment.

2.1   The Games

Our first game is a zero-sum asymmetric matching pennies game introduced by Rosenthal et al. (2003)with the normal form representation given in Figure 1. The game is called Pursue-Evade because the Row player "captures" points from the Column player on a match, and the Column player avoids a loss when he avoids a match. In the game each player can move either Left or Right. 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.

The second game we employ is a unprofitable game introduced by Shachat and Swarthout (2004) referred to as Gamble-Safe. Each player has a Gamble action (Left for each player) from which he receives a payoff of either two or zero and a Safe action (Right for each player) which guarantees a payoff of one. The normal

Column Player

	Left	Right
Left	1 , -1	0 , 0
Right	0 , 0	2 , -2

Row Player (label on left side)

Figure 1: The Pursue-Evade game.

form representation of the game is given in Figure 2. This game has a unique Nash equilibrium in which each player chooses his Left action with probability one-half, and his expected equilibrium payoff is 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 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 Nash equilibrium solution concept than the zero-sum Pursue-Evade game.

Column Player

	Left	Right
Left	2 , 0	0 , 1
Right	1 , 2	1 , 1

Row Player (label on left side)

Figure 2: The Gamble-Safe game.

2.2 Subject pool, recruitment, and general practices

We conducted the experiment sessions in the Economic Science Laboratory at the University of Arizona and the IBM T.J. Watson Research Center. We report results from six sessions, using a total of 60 undergraduate students. Each session contained between 8 and 16 subjects. The subjects were evenly divided between the Pursue-Evade and the Gamble-Safe game treatments. Half of the subjects were assigned Row player roles, and the other half were assigned Column player roles. 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.[2] Then, 200 repetitions of the game were played. For the Pursue-Evade game, Column subjects were initially endowed with a balance of 250 tokens each, and Row subjects none. Each token was worth ten cents. Each subject's total earnings consisted of either a $5 (University of Arizona) or a $15 (T.J. Watson Research Center) show-up payment plus the monetary value of his token balance after the 200th repetition. While a mathematical possibility, no Column subjects went bankrupt.

2.3 Experimental procedures

At the beginning of each repetition, a 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, a 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 payoffs from the specific repetition.

In each session, a Row subject and a Column subject played against each other for the first twenty-three repetitions of the game. Then, beginning in repetition 24, paired subjects stopped playing against each other and for the remainder of the experiment played against a computerized algorithm that implemented the Nash equilibrium strategy.[3] We made the difficult choice to conceal this transition from the subjects. Our motivation was to observe human behavior conditional on their beliefs about how to infer another human's strategy choice and preferences. Further, we were concerned that people will behave differently when they are informed that their opponent (or more generally another beneficiary of their action) is not another person, as was found in prior studies (Eckel and Grossman, 1996; McCabe, Houser, Ryan, Smith, and Trouard, 2001; Fehr and Tyran, 2007).

3 Data summaries and hypothesis tests

A informative starting point is to inspect each subject's action choice proportions versus those of his computerized Nash equilibrium opponent. We present this view of the data for those subjects who were Pursue-Evade

---

[2]Screenshots of the instructions are available at `http://www.excen.gsu.edu/swarthout/mse`

[3]This initial phase of human versus human play is a byproduct of the fact that these data were originally collected as part of another study presented in Shachat and Swarthout (2008) From period 24 onward, computers generated the opponents' action choices. However, these actions were not revealed until both subjects had selected their actions, thereby preserving the pacing of decisions of the human pair.

Row players in Figure 3, Pursue-Evade Column players in Figure 4, Gamble-Safe Row Players in Figure 5, and Gamble-Safe Column players in Figure 6. 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. Within each of these figures is a collection of arrows. Each arrow is a summary of play for a single human-computer pair. The origin of the arrow is located at the joint frequency of Left play in stage games twenty-four through one hundred, and the tip of the arrowhead is located at the joint frequency of Left play in the final one hundred stage games. These arrows show the adjustments subjects make from the first half of stage games to the second half.
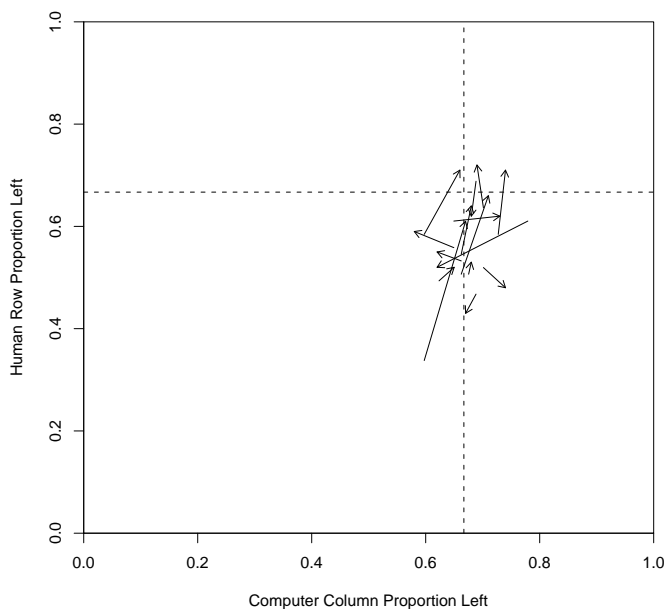

Figure 3: Pursue-Evade joint Left frequencies of human Row players vs. computer NE. Each arrow represents a subject-computer pair. The tail of the arrow is the joint frequency of Left play in rounds 24-100, and the head is the joint frequency of Left play in rounds 101-200.

In Figure 3, we see in the Pursue-Evade game that the Human Row subjects' frequency of Left play is contained within the fifty to seventy percent range and its difficult to ascertain whether Nash equilibrium or equi-probable play better describes the data. For the Human Column subjects in the same game, Figure 4 shows greater heterogeneity; there is a cluster of observations around 50 percent and also some subjects who play Left or Right nearly exclusively. In Figure 5, we see that Human Row subjects in the Gamble-Safe game appear on average to play the Gamble action less than fifty percent, but the range of left frequencies is quite broad covering a range of zero to seventy percent. For the Human Column subjects in this game, Figure 6 suggests that there is a bias towards the minimax strategy as almost all frequencies of Left are below one-half. We will now evaluate our hypotheses regarding specific alternative strategies.
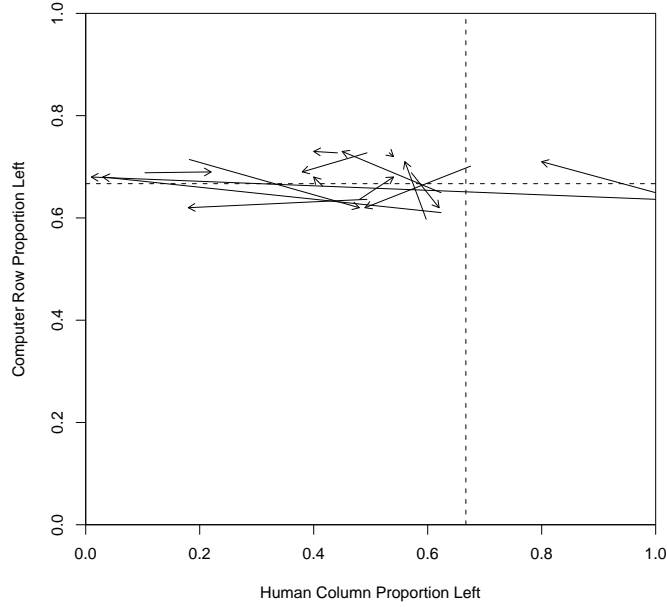
Figure 4: Pursue-Evade joint Left frequencies of human Column players vs. computer NE. Each arrow represents a subject-computer pair. The tail of the arrow is the joint frequency of Left play in rounds 24-100, and the head is the joint frequency of Left play in rounds 101-200.
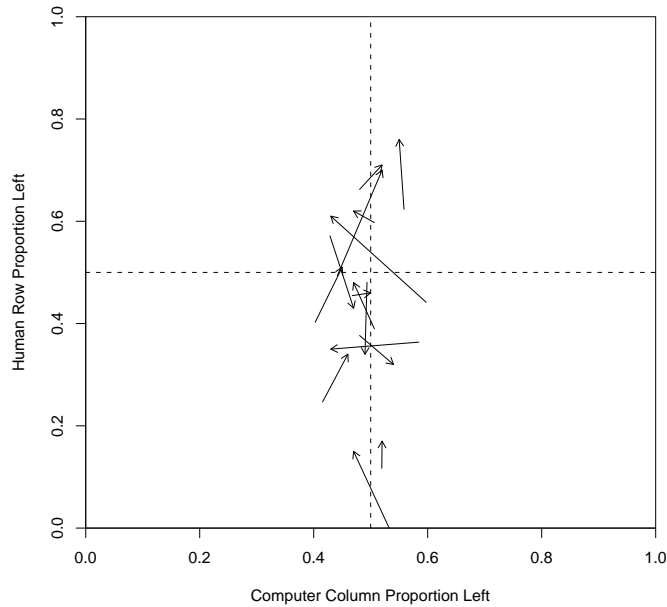

Figure 5: Gamble-Safe joint Left frequencies of human Row players vs. computer NE. Each arrow represents a subject-computer pair. The tail of the arrow is the joint frequency of Left play in rounds 24-100, and the head is the joint frequency of Left play in rounds 101-200.

Hypothesis 1 *Subjects play the Nash equilibrium strategy.*

We first test this hypothesis in the aggregate by pooling the last one hundred rounds within each of the games and player roles, and then conducting a z-test that the proportion of Left play is sixty-seven percent
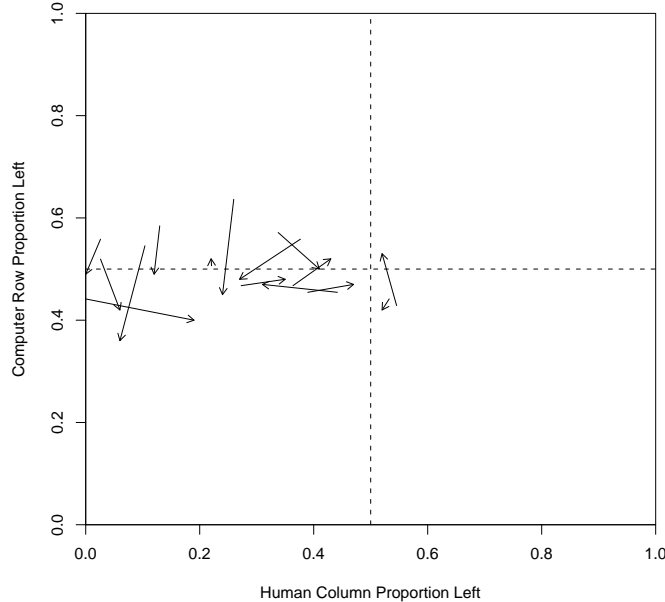
Figure 6: Gamble-Safe joint Left frequencies of human Column players vs. computer NE. Each arrow represents a subject-computer pair. The tail of the arrow is the joint frequency of Left play in rounds 24-100, and the head is the joint frequency of Left play in rounds 101-200.

for the Pursue-Evade sessions, and fifty percent for Gamble-Safe sessions. Table 1 presents the results of these hypothesis tests. The results for the Gamble-Safe game are found in the last two columns of the first row, and the results for the Pursue-Evade game are found in the first two numerical columns of the second row. The Nash equilibrium is strongly rejected in each case. However, the figures of joint play reviewed above suggest a fair amount of heterogeneity in subject play, which would permit the possibility that only some subjects adapt their Nash strategy. To test the hypothesis of homogeneity, subjects use a common mixed strategy, versus the alternative of heterogeneity we conduct a χ^2 test with the results given in the third row of Table 1. Except in the case of Row Pursue-Evade subjects, we reject homogeneity of play. This suggests we should test the Nash equilibrium hypotheses separately for each individual subject.

Next we use the binomial test to determine whether each subject is playing his Nash frequency of Left. For the Pursue-Evade game, the mixed-strategy equilibrium prediction for Left is two-thirds. We present results in Table 2 with each subject's frequency of Left play during rounds 101 through 200. A two-tailed binomial test at the 95 percent level of confidence gives us critical regions of less than 0.58 and more than 0.76 for the final one hundred rounds of play. We reject the Nash proportion of Left play for six out of fifteen Row subjects and for fourteen out of fifteen Column subjects. These results are also summarized in the last row of Table 1.

We conduct a similar set of tests for the Gamble-Safe game where the Nash proportion is one-half. In this case, the critical regions of the two-sided binomial test for the final one hundred rounds are less than

9

Table 1: Aggregate hypothesis tests on final 100 rounds

Null Hypothesis	Test	P-E Row	P-E Col	G-S Row	G-S Col
Aggregated mixed strategy of Left with Probability of 50%	two-tailed Z-test (p-value)	7.28 (0.00)	-7.23 (0.00)	-2.84 (0.00)	-17.20 (0.00)
Aggregated mixed strategy of Left with Probability of 67%	two-tailed Z-test (p-value)	-5.99 (0.00)	-21.42 (0.00)	–	–
All subjects use same mixed strategy	χ^2 test (p-value)	18.14 (0.20)	162.23 (0.00)	108.46 (0.00)	147.71 (0.00)
Individual subject's mixed strategy is 50%	binomal test #rejections at 95% level of confidence	8	7	11	10
Individual subject's mixed strategy is less than 50%	binomal test #rejections at 95% level of confidence	–	–	6	10
Individual subject's mixed strategy is 67%	binomal test #rejections at 95% level of confidence	6	14	–	–

Table 2: Pursue-Evade individual subject hypothesis tests on final 100 rounds.

	Row Player			Column Player		
		Runs test			Runs test	
Pair	% Left	statistic	p-value	% Left	statistic	p-value
1	0.62^e	-2.59	0.01^i	0.40^n	-1.26	0.21
2	0.43^n	-5.13	0.00^i	0.54^n	-0.95	0.34
3	0.66^e	-2.22	0.03^i	$0.18^{n,e}$	-3.26	0.00^i
4	0.48^n	-5.42	0.00^i	0.56^n	-2.50	0.01^i
5	0.61^e	-9.63	0.00^i	0.62^e	1.68	0.09
6	0.71^e	-1.02	0.31	0.45^n	-3.35	0.00^i
7	0.53^n	-0.97	0.33	0.48^n	-4.21	0.00^i
8	0.59	-0.08	0.94	$0.03^{n,e}$	-5.30	0.00^i
9	0.55^n	-0.30	0.76	0.54^n	0.67	0.50
10	0.52^n	0.02	0.99	0.40^n	-3.14	0.00^i
11	0.62^e	-3.44	0.00^i	$0.01^{n,e}$	0.14	0.89
12	0.64^e	-3.95	0.00^i	0.49^n	-1.81	0.07
13	0.72^e	-1.83	0.07	$0.22^{n,e}$	-8.04	0.00^i
14	0.52^n	0.02	0.99	$0.38^{n,e}$	1.90	0.06
15	0.71^e	-2.98	0.00^i	$0.80^{n,e}$	-0.63	0.53

[n] Two-sided binomial test rejection of the Nash equilibrium 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.

forty percent and more than sixty percent. The results of these tests are given in Table 3 and summarized in row four of Table 1. Here the Nash hypothesis is rejected for eleven out of fifteen Row subjects and ten out of fifteen Column Subjects. In total forty-one of our sixty subjects behave in a manner inconsistent with the Nash predicted proportions of Left play.

Another implication of mixed strategy Nash Equilibrium play is that the subjects' choices are generated from time-independent draws from a fixed distribution. For each of our sixty subjects we conduct a non-parametric runs test for serial independence in the last one hundred plays of the game. We reject serial independence for twenty-six of the sixty subjects. These test results are given in Tables 2 and 3. Note that twenty-two of these rejections come from a negative test statistic which is below the rejection threshold. Interestingly, these rejections come from too few runs, evidence of positive serial correlation rather than negative serial correlation.

Table 3: Gamble-Safe individual subject hypothesis tests on final 100 rounds. Missing values due to inapplicability of test on data with zero variation.

| | Row Player | | | Column Player | | |
| | | Runs test | | | Runs test | |
Pair	% Left	statistic	p-value	% Left	statistic	p-value
1	$0.34^{n,m}$	-1.54	0.12	$0.12^{n,m}$	-0.06	0.95
2	0.76^{n}	-0.69	0.49	$0.22^{n,m}$	0.49	0.62
3	0.71^{n}	-1.27	0.21	$0.35^{n,m}$	-3.21	0.00^{i}
4	0.61^{n}	-2.02	0.04^{i}	$0.06^{n,m}$	0.67	0.51
5	$0.15^{n,m}$	-7.36	0.00^{i}	$0.24^{n,m}$	0.14	0.89
6	0.62^{n}	0.19	0.85	0.52	0.22	0.83
7	0.51	-1.00	0.32	$0.00^{n,m}$	$-$	$-$
8	0.48	-3.00	0.00^{i}	$0.06^{n,m}$	0.67	0.51
9	$0.34^{n,m}$	0.92	0.36	0.43	-2.05	0.04^{i}
10	0.46	2.69	0.01^{i}	0.52	0.22	0.83
11	$0.32^{n,m}$	1.04	0.30	0.41	1.79	0.07^{i}
12	$0.35^{n,m}$	0.55	0.58	0.47	0.64	0.52
13	0.70^{n}	-2.40	0.02^{i}	$0.27^{n,m}$	2.45	0.01^{i}
14	0.43	-2.88	0.00^{i}	$0.31^{n,m}$	-1.36	0.17
15	$0.17^{n,m}$	-2.95	0.00^{i}	$0.19^{n,m}$	-1.24	0.21

[n] Two-sided binomial test rejection of the equiprobable Nash equilibrium proportion of 2/3 at the 5% level of significance.
[m] One-sided binomial test rejection of equiprobable proportion at the 5% level of significance in favor of the alternative that play is less than 50%.
[i] Runs test rejection of serial independence at the 5% level of significance.

Equiprobable play, or playing left with a fifty percent probability, play a reasonable alternative to Nash Equilibrium in this setting as the expected payoff is the same for both actions. Of course, for the Gamble-Safe game, the Nash equilibrium strategy and equiprobable play both yield the same prediction. However, in the Pursue-Evade game, Nash equilibrium is distinct from equal probable play. For those subjects who play the Pursue-Evade game, we will test the following hypothesis:

Hypothesis 2 *Subjects play the equiprobable mixed strategy.*

To test for equiprobable play in the Pursue-Evade game, we pool the data for the last one hundred periods for each Row and Column subject, and see that a z-test rejects the hypothesis that the proportion of Left and Right play are equal (see the first row of Table 1). In Table 2 we also conduct a two-sided

binomial test on each subject's choices at the ninety-five percent level of confidence to evaluate individual equiprobable play, and we reject eight out of fifteen Row subjects and seven out of fifteen Column subjects. We conclude that one-half of the Pursue-Evade subjects are behaving consistent with equal-probable play. Next, we consider the minimax model of behavior as an alternative to mixed-strategy equilibrium play in the Gamble-Safe game.

Hypothesis 3 *Subjects play the pure minimax strategy in the Gamble-Safe game.*

For the Gamble-Safe game, the minimax strategy Right is the natural alternative to Nash Play. For each subject we conduct a one-tailed binomial test at the 95 percent level of confidence for each subject's last one hundred decisions to see whether Left is played statistically significantly less than fifty percent. While this is not the most extreme test of minimax play, we nonetheless view it as an appropriate threshold test to determine whether play deviates away from mixed strategy equilibrium play in the direction of minimax play. The critical region for this test is less than thirty-eight. The results of the test are reported Table 3, and summarized in row five of Table 1. For six of the Row subjects and ten of the Column subjects we reject Nash play of fifty percent Left in favor of the alternative of less than 50 percent. So minimax does attract play but not exclusively.

4 A model of strategy heterogeneity and switching

Three factors that confounding the analysis of experimental game data are the latency of subject's 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 and estimate a model that accommodates these idiosyncracies. The results provide a characterization of the heterogeneity, strategy choice dynamics, and estimation of the limiting distribution of mixed strategies employed by our subjects. This distribution has people spending significant time following pure strategies and the equiprobable strategy. We also see there is little self enforcing power of the mixed strategy equilibrium unless it coincides with the equiprobable strategy. Furthermore we uncover several interesting role effects in our two games.

4.1 Hidden Markov model: definition and inference

Consider the following hidden Markov model (HMM hereafter) for a fixed game and player role.[4] The state space S is an n-element subset of subject i's possible mixed strategies. Denote $s_{i,t} \in S$ for the strategy used

---

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

by subject i in period t, S_i is the set of all possible 177 element sequences[5] of mixed strategies for i with typical element s_i, and let $\mathbf{s}$ be the collection of s_i for all 15 subjects. Let $y_{i,t}$ denote player i's realized action in period t, y_i is the corresponding 177 element sequence of i's observable actions, and $\mathbf{y}$ is the collection of y_i for all 15 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. 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 $t - 1$.

For the Pursue-Evade game, we restrict S to contain four elements. When we treat B as fixed and known, the four elements are the pure Left strategy, the Nash equilibrium strategy of two-thirds, the equiprobable mixed strategy, and the pure Right strategy. To assess the validity of this restriction, we subsequently assume B is unknown and estimate the values of its components, and then gauge whether these are close to the values assigned when we treat B as fixed. We make similar formulations for the the Gamble-Safe game, except we specify S as having only three elements. When we treat B as fixed, its elements are the pure Left strategy, the Nash equilibrium (or equiprobable) mixed strategy, and the pure Right (or minimax) strategy. We can now express the likelihood function of (B, π, P) as

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

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 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

---

[5]We only model the second phase of the experiment in which the subjects play against a computerized agent who plays its part of the Nash equilibrium. In this section, we refer to periods 1 through 177, even though in actuality these are the experimental rounds 24 through 200.

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

$$L(B, \pi, P|\mathbf{y}, \mathbf{s}) = \prod_{i=1}^{15} \sum_{\mathbf{s} \in \mathbf{S}} \pi(s_{i1}) B_{s_{i1}}^{I\langle y_{i1}\rangle} (1 - Bs_{i1})^{1-I\langle y_{i1}\rangle} \prod_{t=2}^{177} P_{s_{it-1}, s_{it}} Bs_{it}^{I\langle y_{it}\rangle} (1 - Bs_{it})^{1-I\langle y_{it}\rangle},$$

where $I\langle y_{it}\rangle$ is an indicator function which equals one for the action Left and zero for the action Right. Evaluating this likelihood directly is not computationally practical, as the number of calculations is on the order of $400 \cdot n^{200}$! Thus we employ recursive algorithms, explained shortly, that greatly reduce the computational demands in evaluating this likelihood.

At this point we could proceed down a frequentist's path of directly maximizing the the likelihood function using some variation of the *EM* (expected maximum likelihood) algorithm. Instead we use a Bayesian approach. This choice was pragmatic. In this application the *EM* algorithm converges to different local maximum for different initial parameter values. However, the Bayesian procedure we adopt always converges to the same values regardless of initial values, and these values correspond to the highest of the local maxima found from the *EM* estimation.

In the Bayesian analysis, we first factor the posterior distribution of the unknown HMM parameters and unobserved states $\mathbf{s}$ into the product of posterior conditional distributions. Then we evaluate these 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. We then use these to make inferences regarding posterior distributions 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 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 $(B^{(l)}, P^{(l)}, \pi^{(l)})$ 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 $(B^{(l-1)}, P^{(l-1)}, \pi^{(l-1)})$. This procedure will avoid the large number of calculations needed to evaluate $L(B, \pi, P | \mathbf{y}, \mathbf{s})$. 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 $s_{i, \neq t}$ and other information, $s_{i,t}$ can assume three (Gamble-Safe game) or four (Pursue-Evade game) possibilities, and its conditional posterior distribution is

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

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

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

and for $s_{i,1} = k$

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

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_{ij} and π depends 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 $h(p_{j1}, \ldots, p_{jn}|\eta_{j1}, \ldots, \eta_{jn})$. We have a prior belief that the probabilities of persistence are larger than the probabilities of transition, accordingly we set $\eta_{jj} = 2$ and $\eta_{jk} = 1$ with $j \neq k$. In this study 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 posteriors of $\pi^{(l)}$ 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 $\nu_0 j$ is the number incidences of $s_{i1}^{(l)} = j$ 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

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

and

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

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

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

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

Step 3: Sampling the state dependent mixed strategies B

In our main analysis, we assume B corresponds to the pure strategies Left and Right, the Nash equilibrium strategy, and the equiprobable strategy when it differs from the Nash strategy. 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 make, 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 its 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 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. 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.

4.2 Results of the Bayesian statistical analysis

We begin with analysis of the Pursue-Evade game. First we report the means and variances of the posterior distributions of the transition probability matrix and the initial distribution over states. Then we examine

how well the HMM tracks the dynamics of action choice proportions, and follow that with a description of estimated dynamics of the latent mixed strategy heterogeneity. Finally, we provide an assessment of the robustness of some of our assumed priors.

For each of the Row and Column subject data sets from the Pursue-Evade game, we run the Gibbs sampler for forty thousand iterations. Using the last ten thousand iterations we establish that the empirical density functions have converged via diagnosis by the Geweke test (Geweke, 1991). Then we use the last ten thousand iterations to make statistical inferences. Table 4 presents the estimated means and standard deviations (in parentheses) of the transitional 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.

Recall that our analysis begins with round twenty-four choices after the conclusion of human-human interaction. Thus our estimation of the initial distribution over states, is actually an estimate of the distribution of strategies adopted after twenty-three periods of human play. Inspection of the estimated π's for Row and Column, in the fifth numeric row of Table 4, suggests that human-human interaction concluded without having converged to Nash equilibrium. In both cases, the estimated π places significant probability on each of the four considered mixed strategies. Further less than eighteen and twenty-six percent of the Column and Row players, respectively, are playing the Nash equilibrium strategy.

Next, we discuss the estimated transition matrices P. Inspection of the estimated P's, in the first four numeric rows of Table 4 reveal surprising inertia in the strategies subjects adopt. This is seen from the large values on the main diagonals and corresponding small values on the off-diagonals. When there is a transition between strategies, one mixed strategy is more likely to switch to the other mixed strategy, while there isn't a strong pattern to where pure strategies switch. There are some player role differences worth noting. Column players have a higher probability of persisting in pure strategy play. Also, for the Column player the probability of continuing in the Nash equilibrium strategy conditional upon residing in the Nash state is lower than continuation of the other three strategies.

We also report the long run properties of the estimated Markov chains by calculating the limiting distribution of each estimated transition probability matrix. The limiting distrubtions for both players have the equiprobable model as the modal state, approximately forty percent and fifty-four percent for Row and Column respectively. For the Row player, we also see the Nash strategy receives almost one-third of the limiting distribution probability and each of the pure strategies between twelve and fifteen percent of the probability. One the other hand, for the Column player, pure Right receives almost one-forth of the probability, while the Nash strategy receives the least with less than ten percent.

Next we consider how well the estimated HMM coincides with the observed proportion of Left play over time. For periods $t = 1, \ldots .177$, we calculate the predicted proportion of Left play by the Column (Row)

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

	Row Player				Column Player			
	PR	EM	NE	PL	PR	EM	NE	PL
PR	0.867 (0.026)	0.038 (0.018)	0.043 (0.019)	0.052 (0.022)	0.923 (0.013)	0.026 (0.011)	0.034 (0.014)	0.017 (0.007)
EM	0.009 (0.004)	0.909 (0.028)	0.058 (0.026)	0.025 (0.011)	0.011 (0.005)	0.996 (0.012)	0.015 (0.009)	0.007 (0.003)
NE	0.015 (0.008)	0.066 (0.031)	0.878 (0.042)	0.040 (0.024)	0.076 (0.040)	0.086 (0.038)	0.788 (0.061)	0.049 (0.036)
PL	0.056 (0.020)	0.073 (0.031)	0.073 (0.029)	0.798 (0.034)	0.36 (0.014)	0.03 (0.013)	0.037 (0.015)	0.898 (0.019)
π	0.115 (0.085)	0.621 (0.182)	0.178 (0.158)	0.086 (0.077)	0.066 (0.062)	0.376 (0.184)	0.262 (0.191)	0.296 (0.131)
Limiting Distribution	0.125	0.405	0.324	0.146	0.234	0.540	0.098	0.128

Note: Standard deviations in parentheses.

subjects, $\widehat{Left}_t$, by

$$\widehat{Left}_t = \sum_{l=30000}^{40000} \sum_{d=1}^{15} \sum_{j=1}^{4} (I_{s_{d,t}^l = j}) \cdot B_j / (15 \cdot 10000).$$

Figures 7 and 8 provide evidence these predicted proportions of Left play track the actual proportions quite well. Admittedly this is an in-sample forecasting exercise, but at the same still impressive, as minimizing forecast error is not the objective of our statistical inference exercise.
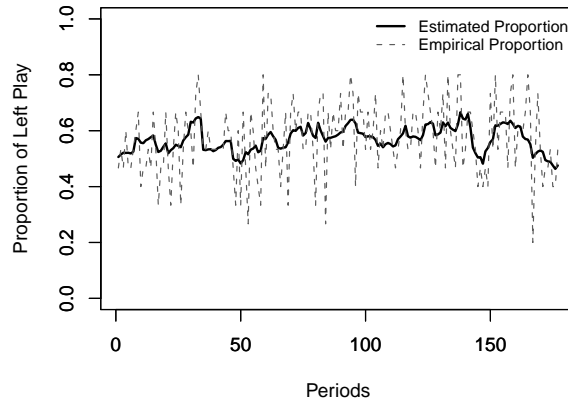

Figure 7: Proportion of Row player's Left choice over periods in Pursue-Evade game

Figures 9 and 10 each present times series the estimated proportion of subjects using each of the four strategies evolves over time. For strategy j the estimated proportion of subjects using that strategy in a given round t is

$$\hat{j}_t = \sum_{l=30000}^{40000} \sum_{d=1}^{15} (I_{s_{d,t}^l = j}) / (15 \cdot 10000).$$

For the Row players, the initial proportion of equiprobable strategy players is two-thirds which steadily
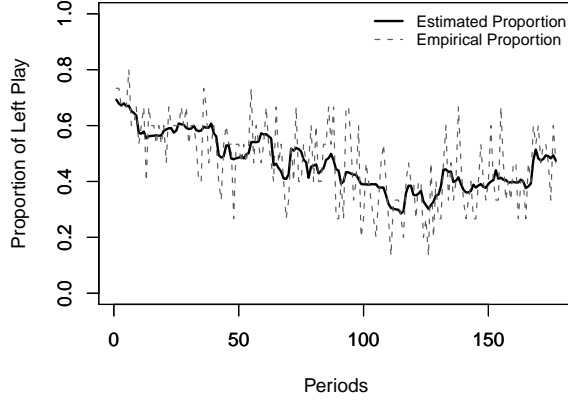
Figure 8: Proportion of Column player's Left choice over periods in Pursue-Evade game

declines in the first fifty rounds of human-computer interaction. Meanwhile, we find some small evidence of self-enforcement of the Nash equilibrium, as the proportion of Nash strategy adopters rises slightly. We also observe that the pure Left and Right strategies states consistently remain below twenty percent. The results for the Column subjects are quite different. The use of the Nash strategy steadily declines while the adoption of the equiprobable strategy rises. Furthermore,we see a strong emergence of pure Right play, and conversely a steady decline in pure Left play.
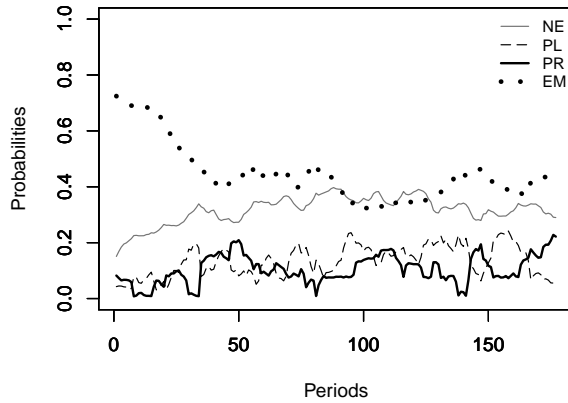
Figure 9: Strategy dynamic of Row player over periods in Pursue-Evade game

Next we provide two different robustness evaluations of our Bayesian analysis of the HMM. First, we assess the appropriateness of our degenerate prior on B by redoing the Gibbs sampler exercise 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 Figures 11 and 12 we present kernel smoothed plots of these approximations to posterior densities. Inspection reveals these four posteriors are sharply peaked and closely centered on our assumed four strategies.
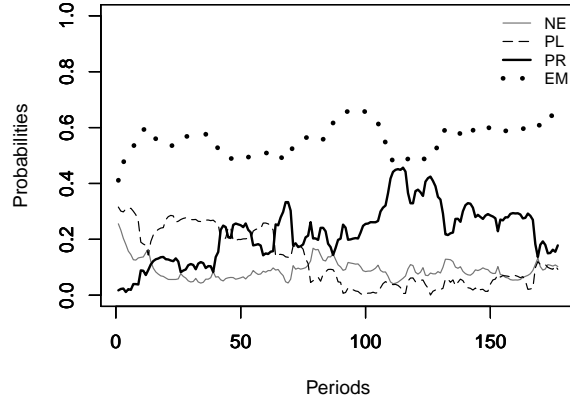
20

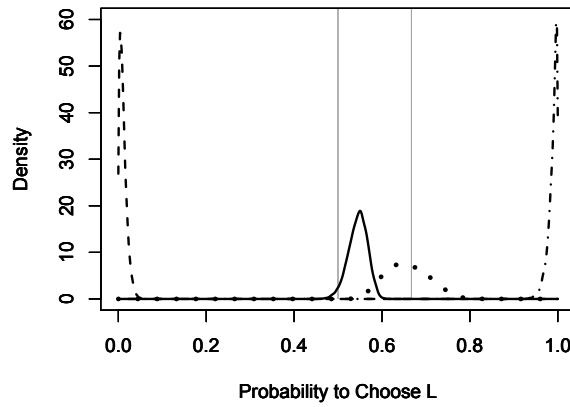Figure 10: Strategy dynamic of Column player over periods in Pursue-Evade game

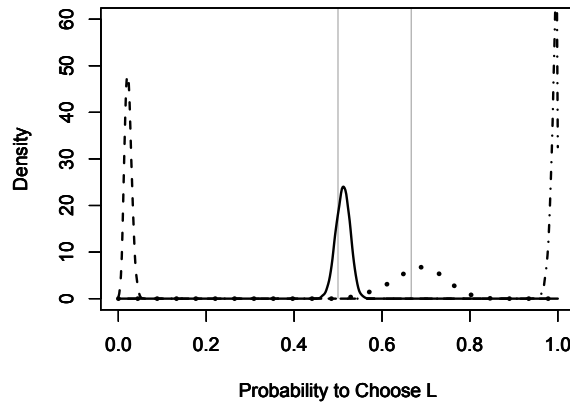Figure 11: Posterior distribution of B for Row player in Pursue-Evade game

Figure 12: Posterior distribution of B for Column player in Pursue-Evade game

Next we assess the assumed heterogeneity and dynamic structure of the model. First we calculate log-likelihoods assuming all subjects follow in turn the Nash equilibrium mixed strategy, the equiprobable mixed strategy, and then a pure mixture model in which each subject is assumed to use the same strategy the

entire sample but this strategy can vary across subjects.[6] We compare each of these with the HMM by way of log-likelihood ratio tests reported in Table 5. We see that the HMM consistently provides statistically significantly better explanation of the data, with test p-values essentially zero in every case.

Table 5: Log-likelihood ratio tests of alternative models in the Pursue-Evade game

	HMM		NE Strategy		EM Strategy		Mixture Model	
	Row	Column	Row	Column	Row	Column	Row	Column
−LL	1356	1183	1863	2055	1840	1840	1796	1802
LLR Test	–	–	1014	1744	969	1316	881	1238
(p-value)	–	–	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

We now turn our attention to the Gamble-Safe game. For both the Row and Column player data sets we ran the Gibbs Sampler for twenty thousand iterations, using the last ten thousand iterations for inference after testing for convergence of the empirical densities via the Geweke test. The posterior means and standard deviations are reported in Table 6. First consider the results for the Row player. Here the Markov transition matrix almost fails irreducibility (roughly meaning we can always reach one state from another, even if it takes more than one transition.) The probability of continuing in the Nash equilibrium state is nearly one, and the two pure strategy almost always transition between each other. Interestingly the pure Left state transitions to the pure Right state with a probability of over sixty percent. Correspondingly the limiting distribution places over ninety percent probability on the Nash strategy and about eight percent on the pure Right strategy. Of course, this is agreeable with the two main game theoretic solutions we have for this game; Nash equilibrium and minimax.

Table 6: Estimated transition matrices and limiting distributions of Gamble-Safe game

	Row			Column		
	PR	NE	PL	PR	NE	PL
PR	0.896 (0.016)	0.009 (0.007)	0.095 (0.016)	0.876 (0.010)	0.005 (0.003)	0.119 (0.010)
NE	0.001 (0.001)	0.998 (0.001)	0.001 (0.001)	0.003 (0.003)	0.993 (0.003)	0.004 (0.003)
PL	0.489 (0.066)	0.068 (0.044)	0.443 (0.063)	0.710 (0.035)	0.024 (0.017)	0.266 (0.033)
π	0.374 (0.122)	0.493 (0.138)	0.132 (0.096)	0.441 (0.121)	0.376 (0.131)	0.183 (0.102)
Limiting Distribution	0.082	0.902	0.016	0.406	0.525	0.069

Note: Standard deviations in parentheses.

The posterior mean of the transition matrix for the Column role has similar features. The minimax strategy also has a continuation probability close to one. However, the Left strategy state now transitions

---

[6]Since no one subject plays only Left or only Right the last 177 periods, these strategies are not included as they have a likelihood value of zero.

to the Right strategy state more than seventy percent of the time. This results in a limiting distribution quite different from that of Column. Now we see fifty-two percent of the limiting probability is on the Nash strategy and forty percent on the minimax strategy. The minimax strategy appears to be a more oft choice for the Column rather than the Row subjects.

Next we present in Figures 13 and 14 the time series of predicted and observed proportions of Left play for Row and Column subjects, respectively, in the Gamble-Safe game. For the Row player, the observed proportions are rather flat over time and the predicted values track this closely. There is a more interesting slight U-shaped trend for the Column role, which the predicted values mimic.
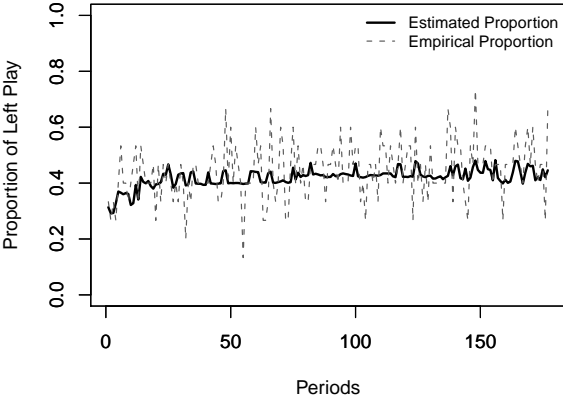

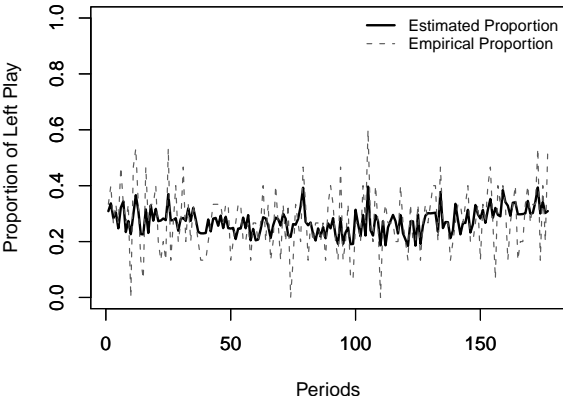Figure 13: Proportion of Row player's Left choice over periods in Gamble-Safe game

Figure 14: Proportion of Column player's Left choice over periods in Gamble-Safe game

Finally, we consider the estimated dynamic latent strategy profiles for the Gamble-Safe game. 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. First, consider the Row players in Figure 15. We see a steady smooth growing trend of the Nash strategy. In contrast, we see a jagged mirror image of the Left and Right strategy series. Looking closely at those two series, we can

see the clear negative correlation generated by the cycling between the Pure strategies. The Column graph shown in Figure 16 has an even more interesting set of series. Once again, the proportion of Nash states has a smooth, but this time non-monotonic, adjustment path. Meanwhile we once again see that the Left and Right strategy state proportions have the jagged mirror relationship characteristic of cycling.
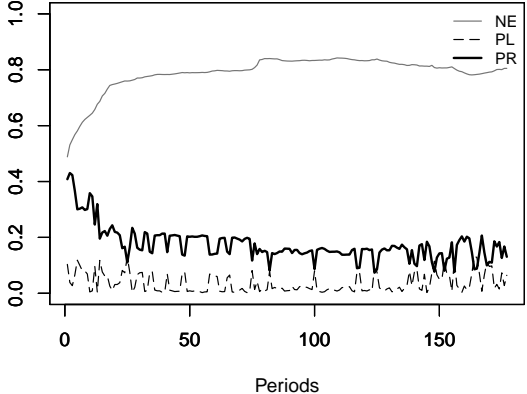

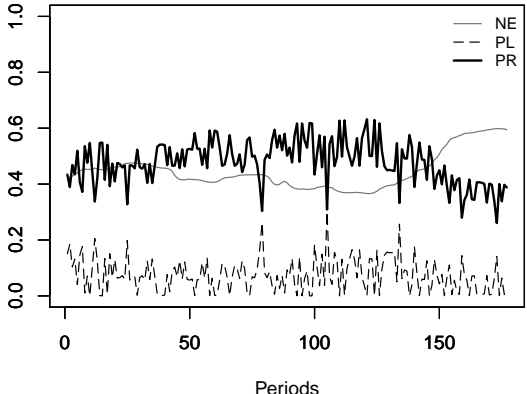Figure 15: Strategy dynamic of Row player over periods in Gamble-Safe game

Figure 16: Strategy dynamic of Column player over periods in Pursue-Evade game

We provide robustness checks for the Gamble-Safe HMM inference as we did for the Pursue-Evade game. In this case we only consider one strictly mixed strategy, as this is the only hypothesized strategy that doesn't suffer from a zero-likelihood problem when making a log-likelihood comparison to the HMM. The differences in the log likelihoods are staggering: 1244 and 1604 for the Row and Column roles, respectively (which have p-values of virtually zero in a likelihood ratio hypothesis tests). We also test the robustness of our strong restriction that $B = [0, 0.5, 1]$, by considering an HMM where these state conditional strategies each have a uniform Beta prior. The kernel smoothed empirical densities of the posteriors are presented in Figures 17 and 18 for Row and Column players, respectively. The posterior for the Row players is consistent with our strong prior. Unfortunately, the same can not be said for the Column players. In this case, one

24

state conditional strategy is clearly below one-half and the other has its mass in the interval 0.03 to 0.08. Overall, this is the only really weak diagnostic result of the HMM inferences. And the methodology has proven itself an effective tool at modeling the heterogeneity of latent mixed (or not mixed) strategy play and the dynamics of strategy evolution.
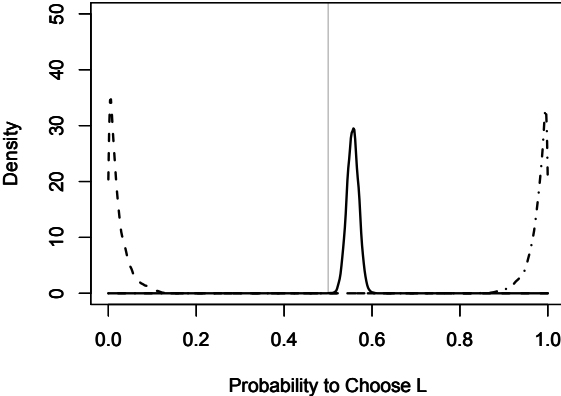

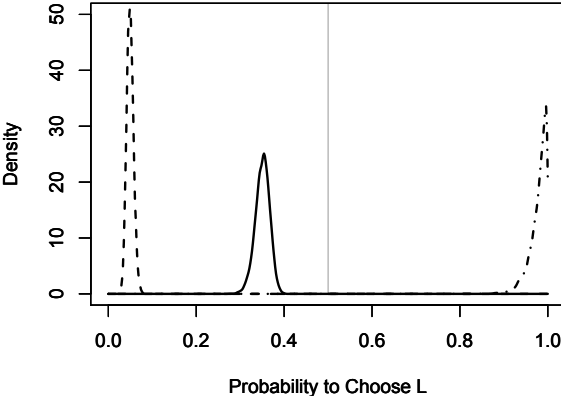Figure 17: Posterior distribution of B for Row player in Gamble-Safe game

Figure 18: Posterior distribution of B for Column player in Gamble-Safe game

5 Conclusions

In this study we have examined the self-enforcing nature of a unique Nash Equilibrium in mixed strategies. Our experiment is particularly well suited for this purpose, as each subject plays against a computer which generates actions according to its mixed strategy equilibrium strategy. A consequence of this experimental design is that any failure of a human subject to play the Nash strategy does not result from the non-convergence of a learning dynamic but rather from the lack of self enforcement. Thus, disentangling two effects which have been difficult to identify in previous studies. Furthermore, our study is another example

for the growing literature which studies human and computer agent interactions in important economic, organization, and social settings. Understanding these types of interactions is of upmost importance as these interactions are becoming increasingly commonplace in our daily lives.

In terms of the self-enforcement of strictly mixed Nash equilibrium, we consider two games. Pursue-Evade, a zero-sum game, provides the self-enforcement hypothesis its best chance: the Nash equilibrium and minimax solutions coincide, the impact of other-regarding preferences on the equilibrium is minimized, and there no room to develop cooperation through repeated-game strategies. The second game, Gamble-Safe, is chosen because it presents a strong challenge to self-enforcement. The game is unprofitable, meaning that the minimax and Nash equilibrium strategies differ, but offer the same expected game payoff. Furthermore, in this game the minimax strategy is pure, which we conjecture provides additional attraction to this alternative solution.

In the constant-sum Pursue-Evade game we reject Nash equilibrium play for half of the Row subjects and all but one of the Column subjects. We also reject equiprobable play for about half of both Row and Column subjects. Moreover, we find that despite all subjects playing against the same equilibrium strategy, there is significant heterogeneity across subjects. In the unprofitable Gamble-Safe game, we reject the Nash equilibrium (and equiprobable) model for two-thirds of both the Row and Column players. Further, we see many of the human Column players selecting Left well below fifty percent of the time, indicating a tendency towards the minimax safe strategy. However, some of these hypotheses tests are based on the assumptions of homogeneity of strategy choice across subjects and all assume time invariance of this choice. Since evidence suggests both of these assumptions are violated, we introduce a hidden Markov model that allows for these violations.

The hidden Markov model treats the players' mixed strategies as latent data and specifies a first order Markov dynamic for strategy adjustment. This model proves quite useful in modeling this heterogeneity and is a new way to estimate whether subjects play according to mixed or pure strategies. Analysis of the HMM suggests that the equiprobable model has more attraction than the Nash equilibrium mixture when they do not agree. Furthermore, we find a significant amount of Pure strategy play. We also identify strong cycling patterns between pure strategies, concurring with prior research which directly elicited subjects' mixed strategies (Shachat, 2002). Overall, the HMM approach is quite promising and an obvious application would be to apply it to the learning patterns in human versus human repeated normal form game experiments.

References

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

Aumann, R. J., and Maschler, M. (1972). Some thoughts on the minimax principle. *Management Science*, *18*, 54–63.

Binmore, K., Swierzbinski, J., and Proulx, C. (2001). Does minimax work? an experimental study. *Economic Journal*, *111*, 445–64.

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

Brown, J., and Rosenthal, R. (1990). Testing the minimax hypothesis: A re-examination of O'Neill's game experiment. *Econometrica*, *58*, 1065–1081.

Coricelli, G. (2001). Strategic interaction in iterated zero-sum games. *Technical Report 01–07*, University of Arizona.

Duersch, P., Kolb, A., Oechssler, J., and Schipper, B. C. (2010). Rage against the machines: How subjects learn to play against computers. *Economic Theory*, *43*, 407–430.

Duffy, J. (2006). Agent-based models and human-subject experiments. In L. Tesfatsion, and K. L. Judd (Eds.), *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*, Handbooks in Economics Series, Amsterdam: North-Holland/Elsevier.

Eckel, C. C., and Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, *16*, 181–191.

Erev, I., and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, *88*, 848–881.

Fehr, E., and Tyran, J.-R. (2007). Money illusion and coordination failure. *Games and Economic Behavior*, *58*, 246–268.

Fox, J. (1972). The learning of strategies in a simple, two-person zero-sum game without saddlepoint. *Behavioral Science*, *17*, 300–308.

Geman, S., and Geman, D. (1987). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In M. A. Fischler, and O. Firschein (Eds.), *Readings in computer vision: issues, problems,*

principles, and paradigms, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 564–584.

URL http://portal.acm.org/citation.cfm?id=33517.33564

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

URL http://ideas.repec.org/p/fip/fedmsr/148.html

Harsanyi, J. C. (1973). Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, *2*, 1–23.

Levitt, S. D., List, J. A. L., and Reiley, D. H. (2010). What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments. *Econometrica*, *78*, 1413–1434.

Lieberman, B. (1962). Experimental studies of conflict in some two-person and three-person games. In J. H. Criswell, H. Solomon, and P. Suppes (Eds.), *Mathematical Methods in Small Group Processes*, Stanford University Press, 203–220.

McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 11832–11835.

McKelvey, R. D., Palfrey, T. R., and Weber, R. A. (2000). The effects of payoff magnitude and heterogeneity on behavior in 2 x 2 games with unique mixed strategy equilibria. *Journal of Economic Behavior & Organization*, *42*, 523–548.

Messick, D. M. (1967). Interdependent decision strategies in zero-sum games: A computer-controlled study. *Behavioral Science*, *12*, 33–48.

Mookherjee, D., and Sopher, B. (1994). Learning behavior in an experimental matching pennies game. *Games and Economic Behavior*, *7*, 62–91.

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

Mukherji, A., McCabe, K., and Runkle, D. E. (2000). An experimental study of information and mixed-strategy play in the three-person matching-pennies game. *Economic Theory*, *15*, 421–462.

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

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

Palacios-Huerta, I., and Volij, O. (2008). Experientia docet: Professionals play minimax in laboratory experiments. *Econometrica*, *76*, 71–115.

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

Richiardi, M. G., Leombruni, R., and Contini, B. (2006). Exploring a new expace: The complementarities between experimental economics and agent-based computational economics. *Journal of Social Complexity*, *3*, 1–9.

Rosenthal, R. W., Shachat, J., and Walker, M. (2003). Hide and seek in arizona. *International Journal of Game Theory*, *32*, 273–293.

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

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

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

Shachat, J., and Swarthout, J. T. (2008). Learning about learning in games through experimental control of strategic interdependence. *Technical Report 2008-06*, Experimental Economics Center, Georgia State University.

Spiliopoulos, L. (2008). Humans versus computer algorithms inrepeated mixed strategy games. *Technical Report 6672*, Munich Personal RePEc Archive.

Sussman, A., Tabb, L., and Lati, R. (2009). U.S. equity high frequency trading: Strategies, sizing and market structure. *Technical report*, Tabb Group.

Willinger, M., and Noussair, C. (2003). Efficient mixing and unpredictability in an experimental game, emory University.

Wooders, J. (2010). Does experience teach? professionals and minimax play in the lab. *Econometrica*, *78*, 1143–1154.