.

An Experiment to Evaluate Bayesian Learning of Nash Equilibrium Play

James C. Cox[1], Jason Shachat[2], and Mark Walker[1]

1. Department of Economics, University of Arizona
2. Department of Economics, University of California at San Diego

February 1997
First version: March 1995

Abstract

Some recent theoretical approaches to the question of how players might converge over time to a Nash equilibrium have assumed that the players update their beliefs about other players according to Bayes' Rule. Jordan has shown in a Bayesian model of this kind that play will (theoretically) always converge to a complete-information Nash equilibrium, even though individual players will not generally attain complete information. We report on an experiment designed to evaluate the empirical implications of Jordan's model. A finite version of the model is constructed which generates unique predictions of subjects' choices in nearly all periods. The experimental data reveals that the theory does reasonably well at predicting the equilbria that subjects eventually play, even when there are multiple equilibria. The results thus suggest that Jordan's Bayesian model can provide an empirically effective solution to the equilibrium selection problem when the players have beliefs with finite support. However, the model's predictions about the path of play over time are not consisitent with the experimental data.

We are grateful to the University of Arizona's Economic Science Laboratory (ESL) for research support; to the National Science Foundation for financial support; and to Timothy O'Neill of the ESL for his extraordinary programming services.

An Experiment to Evaluate Bayesian Learning of Nash Equilibrium Play

James C. Cox, Jason Shachat, and Mark Walker

The attempt to rationalize equilibrium analysis in games has recently shifted from arguments based on players' introspection and common knowledge to the idea that equilibrium play is *learned* by the players of a game through repeated play. Some of the contributions to this learning-based approach to equilibrium have the character that the players are fully rational, in the sense that they have well-defined beliefs about one another that are updated in a Bayesian fashion in response to their experience. (See, for example, Jordan (1991) and Kalai & Lehrer (1993).) In this paper we describe an experiment that is designed to shed light on the idea that equilibrium play may be learned by players who behave as Bayesian learners. The experimental results provide some empirical support for the Bayesian learning approach to rationalizing equilibrium analysis.

The experiment is based on the Bayesian learning model analyzed by Jordan (1991). In Jordan's model, players are engaged in a noncooperative normal-form game. Each player knows his own payoff function but he has only a probabilistic belief about the other players' payoff functions. In this game of incomplete information, the players are assumed to play a Bayesian Nash equilibrium, but this will not generally yield a Nash equilibrium of the "true game" they are playing – the game defined by the players' *actual* payoff tables. The question Jordan addresses is whether, by playing the game repeatedly, the players will learn to play a Nash equilibrium of the true game.

Specifically, Jordan assumes that at each stage of play the players play a Bayesian Nash equilibrium based upon their current beliefs, and that after each play has occurred each player updates his belief according to Bayes' Rule by incorporating the play he has just observed. Jordan proves that under certain conditions this process does indeed converge to a Nash equilibrium of the true game. A further interesting feature of Jordan's analysis is that while the players do eventually play a complete-information Nash equilibrium, in most cases they do so without actually becoming completely informed: each player simply learns enough about the others to lead him to play his part of the complete- information Nash equilibrium.

In order to evaluate the empirical implications of Jordan's Bayesian learning theory, we have devised and conducted an experiment based on the instructive example in Jordan's 1991 paper. Jordan's example has a continuum of types (*i.e.*, possible payoff functions) for each player. In designing an experiment, however, it is necessary to limit the players to a relatively small finite set of possible types, and thus to have only a small finite set of possible "true games." Our particular conversion of Jordan's example to one with a small finite set of possible games has provided an interesting collection of 2×2 true games: there are games that have a unique equilibrium which is in pure strategies; a coordination game with opposing interests (essentially, a Battle of the Sexes game); a game in which one equilibrium Pareto dominates the others; and a game with a unique equilibrium in mixed strategies.

Our experimental results indicate that in one important respect the Jordan model's predictions are correct: play does generally move to a complete information Nash equilibrium of the true game the players are playing, even when the players clearly have not attained complete information. Moreover, because of the particular prior beliefs we induce in our subjects, the Jordan theory predicts *which* equilibrium will be attained in the games with multiple equilibria, and the experimental results indicate that play does move toward the predicted equilibrium. For example, perhaps the most interesting result that emerges from the experiment concerns the coordination game. The game is symmetric, so that other theories of learning do not single out either of the pure-strategy equilibria as any more likely to occur than the other. Jordan's theory, however, predicts that under our subjects' induced prior beliefs, a specific one of the equilibria will eventually be played. The equilibrium predicted by the Jordan theory does indeed occur substantially more often in the experiment than the other pure-strategy equilibrium. These results suggest that the players' priors about one another may play a more important role in resolving equilibrium selection problems than has been previously recognized.

Although the experiment provides support for the predictions the Jordan theory makes about eventual play, there is essentially no support in the experimental data for the proposition that play follows the path over time that the Jordan theory predicts. The Jordan theory predicts the exact profile of play in every period for most of the true games that the subjects played, but only about half of the observed play in the first few periods was actually as predicted, with the other half of the plays falling into one of the other three cells of the 2×2 game. This suggests that it would be informative to compare the Jordan model's predictions with those of other models of out-of-equilibrium learning in games. We address this idea in Section 4.

The remainder of the paper is organized as follows. We describe Jordan's model and his example in Section 1. In Section 2 we describe the experiment, as well as some of the considerations behind the specific design we have developed, including the refinement of Jordan's continuum example to one with a small finite set of games. Section 3 contains the procedural details of the experiment. The experimental results are reported in Section 4, and Section 5 presents conclusions and some implications of our results for further research.

1. Jordan s Model of Learning in Games

We provide here a brief description of Jordan's model of Bayesian learning in games, as well as a brief description of the 2×2 example he presents in the same paper. Although Jordan develops the theory for *n*-player games with arbitrary finite strategy sets, we describe the model for two-player 2×2 games (two strategies for each player), which is the framework for the example and for our experiment. Our convention throughout the paper will be to refer to the two players as the Row player and the Column player, with strategy sets {Top, Bottom} and {Left, Right}, or simply {T,B} and {L,R}. We will also refer to the Row player as Player 1 and the Column player as Player 2.

Jordan's model assumes that each player i has a payoff function π_i that gives his payoff at each of the four possible profiles of play that can occur, and has a prior belief about (i.e., a probability distribution over) his opponent's possible payoff functions, or "types." The two players play the game repeatedly, with their given, unchanging payoff functions, in each period playing a Bayesian Nash equilibrium of the stage game for their current probability beliefs about their opponents' types. Thus, the players are assumed to play the game myopically, always attempting to maximize only the current period's payoff. After each play, each player updates his belief according to Bayes' Rule, incorporating the play by his opponent that he has just observed. Jordan refers to a path of play that satisfies the conditions we have just described as a *Bayesian strategy process*, and he shows that for all finite games (except possibly for a set with prior probability zero) every Bayesian strategy process converges to a Nash equilibrium of the "true" game, that is, of the game defined by the players' actual payoff functions.

In his example, Jordan considers all 2×2 games. Each player's set of possible payoff tables, or types, is therefore $\mathbb{R}^4$, but Jordan reduces this to the unit circle by first eliminating all the games in which a player's payoff is completely flat, that is, independent of the strategies chosen, and then by normalizing, as follows. He first distinguishes a strategy's payoffs only by payoff *differences* (which are all that matter for the behavior they imply), so that every game can be represented as having payoff functions of the form shown in Figure 1: the Row player's payoff is always zero if he chooses Bottom, and we can therefore describe his entire payoff function with just the pair (a,b) = (π_1(T,L), π_1(T,R)); and the Column player's payoff is always zero if he chooses Right, so that we can describe his payoff function with the pair (α,β) = (π_2(T,L), π_2(B,L)). Then, since multiplying a payoff function by a positive number also does not change its implied choice behavior, Jordan further restricts the payoff pairs (a,b) and (α,β) to lie on the unit circle. Jordan completes his example by specifying that each player's prior distribution on his opponent's types is the uniform distribution on the unit circle, and then he describes in detail a particular Bayesian strategy process — that is, he describes a path of beliefs and play that can arise from a particular class of payoff-function pairs (π_1, π_2).

2. Experimental Design

In Jordan's example the players' types (their payoff tables) are drawn from a continuum of possible payoff tables. In an experiment, however, the set of types must be finite. Indeed, in the interest of simplicity (in the theory, and for the experimental subjects), the number of types for each player should be kept as small as possible. An additional reason for keeping the number of types small is to minimize the total number of draws (and therefore also time and subject payments) required in order to generate sufficient sample sizes for each possible draw of a type for each player. We have therefore developed a simplified finite analogue of Jordan's example, and in order to further simplify the games for subjects, the games have only integer payoffs. The pairs (a,b) and (α,β) that identify Row and Column payoff functions will of course not lie on the unit circle any longer if they have integer components, but this is immaterial.

Construction of an appropriate set of "induced prior beliefs" for an experiment – i.e. the Row and Column payoff tables, or player types, and the associated probabilities – is not straightforward. It requires a great deal of care, for example, to avoid induced beliefs that generate multiple Bayesian Nash equilibria at every period of play. The theory makes its sharpest predictions, and is thus amenable to a more powerful test, when it makes a unique prediction of behavior in each period of play. Thus, one objective in developing our experimental design was to minimize the number of periods at which multiple Bayesian Nash equilibria can occur and, when several could occur, to minimize their number.

A similar consideration involves mixed-strategy equilibria: if one or both players are predicted to use a mixed strategy in a Bayesian Nash equilibrium, then again the theory's predictions are, at best, difficult to test. Thus an additional objective for our experimental design was to minimize the frequency with which the player types that are randomly drawn will generate mixed-strategy Bayesian Nash equilibrium, and especially to minimize the frequency of drawing types for which the theory predicts convergence to a mixed-strategy Nash equilibrium of the true game.

The following configuration of player types avoids the pitfalls we have just described, and provides a sound design for an experimental analysis of Jordan's model. The design uses the four types,

$$\mathcal{A} = (1,2),\ \mathcal{B} = (-2,1),\ \mathcal{C} = (-1,-2), \text{ and } \mathcal{D} = (2,-1),$$

but incorporates them into different sets for the two players:

$$\mathcal{T}_1 = \{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}\} \text{ and } \mathcal{T}_2 = \{\mathcal{B}, \mathcal{D}\},$$

as depicted in Figures 2a and 2b. We also abandon the uniform distribution for Player 1's prior; the probabilities with which the types will be drawn are (using p for the prior on Player 1's types and q for the prior on Player 2's):

$$p_A = 3/8,\ p_B = 1/8,\ p_C = 3/8,\ p_D = 1/8$$

and

$$q_B = 1/2,\ q_D = 1/2.$$

Table 1 provides a concise summary of the path of expectations and play in the environment $\mathcal{T}_1 \times \mathcal{T}_2$. Note in particular in Table 1 that in every one of the eight games that can be drawn from $\mathcal{T}_1 \times \mathcal{T}_2$ the theory makes a precise, unique prediction about the play that will occur in every period, with the sole exception of periods three and later in the mixed-equilibrium-only games $\mathcal{B}v\mathcal{D}$ and $\mathcal{D}$v$\mathcal{B}$. In period 1 there is a unique pure-strategy Bayesian Nash equilibrium. This Bayesian Nash equilibrium separates Player 2's two types (i.e., the play at period 1 reveals his type), and it separates Player 1's types into the two sets $\{\mathcal{A}, \mathcal{D}\}$ and $\{\mathcal{B}, \mathcal{C}\}$. (Figure 2 should be helpful here.) Then, at period 2, there is again (conditional on each of the four possible plays at $t = 1$) a unique Bayesian Nash equilibrium, which is pure, for each game. This Bayesian Nash equilibrium either reveals Player 1's type (followed by Nash equilibrium play at all subsequent t), or else it is "pooling," in which case Player 1 makes the same play whatever his type (followed by the same behavior at all subsequent periods, since Player 2 learns nothing from it).

Note that in every one of the eight games in $\mathcal{T}_1 \times \mathcal{T}_2$, Player 2's Nash equilibrium play depends upon Player 1's type. Thus, the players can be said to have "learned" to play Nash equilibrium in each of the games, even though in some of the games they play Nash equilibria in every period. Furthermore, in only one-fourth of the draws – the ones in which π_1 is either $\mathcal{B}$ or $\mathcal{D}$ – are there multiple or mixed-strategy Nash equilibria; in the other three-quarters of the draws, in which the Nash equilibrium is unique and pure, it will be clear whether subjects are playing it.

3. Experimental Procedures

In order to eliminate the possibility of subjects losing money and perhaps becoming bankrupt, and to avoid the difficulties that can arise if subjects treat losses and gains asymmetrically, every payoff table in $\mathcal{T}_1$ and $\mathcal{T}_2$ was translated into a table with all payoffs nonnegative, by adding a constant to each of the table's payoffs, as in Figure 3. Thus, the sets $\mathcal{T}_1$ and $\mathcal{T}_2$, as presented to the subjects, were the ones shown in Figure 4. This transformation leaves all the theory for $\mathcal{T}_1 \times \mathcal{T}_2$ unchanged; in particular, Table 1 is unchanged.

The experiment consisted of four sessions of about two hours each. There were 12 subjects in each session, six of whom played the role of Row player throughout, and six of whom always played Column. No subject participated in more than one session.

Jordan's model assumes that players play myopically, always attempting to maximize their current period payoff, and ignoring future payoffs, thereby eliminating such phenomena as reputation building, punishment, and information gathering. In order to minimize the possibility that our experimental subjects would take future play into account, the Row subjects and Column subjects were randomly matched anew with one another in every period, with the added proviso that a given pair of subjects could never be matched with one another in two successive periods. The subjects were informed of the random-matching protocol with non-technical wording in the experiment's instructions.

The experiments were run on the University of Arizona's Economic Science Laboratory networked computer facility. The particular games that the subjects played in a given session were drawn according to the probabilities given in the preceding section (the probabilities also appear in Figure 4). The four payoff tables in $\mathcal{T}_1$ and the two tables in $\mathcal{T}_2$, together with their corresponding probabilities, were all presented on each subject's computer screen and were therefore the prior beliefs about types that were induced.

Each of the four experimental sessions consisted of several "regimes," each of which began with a drawing of a Row payoff table (revealed only to the Row players, by on-screen highlighting of that table) and a Column payoff table (revealed only to the Column players). The resulting game was played for 15 periods, and then a new regime was begun with the drawing of new payoff tables. The four sessions included the regimes presented in Table 2 and Figure 4.

Subjects were paid $5 upon signing in for the experiment. The salient rewards in the games' payoff matrices were converted into U.S. dollar rewards at the rate of six experimental dollars per one U.S. dollar in Session 1 and at a ten to one exchange rate in the other sessions. The salient rewards paid to the subjects are reported in Table 3.

4. Analysis of the Data

We use the data from our experiment to attempt to answer the following four questions:

A. Did the subjects in the experiment learn to play a Nash equilibrium of the true game?

B. In the two true games with multiple Nash equilibria, did the subjects learn to play the unique equilibrium predicted by the Jordan learning model?

C. Is the observed play consistent with the period-by-period predictions that the Jordan model makes about the path of play in our experiment?

D. Does the experimental design adequately discriminate between the Jordan learning model and alternative game-theoretical learning models, such as fictitious play and Cournot best response?

4.1 Learning the Nash Equilibrium

Our experiment is designed so that Jordan's learning model predicts that by period 3, and in all subsequent periods, all play will be at a Nash equilibrium of the true game. The consistency of the data with this prediction is presented in Table 4, which reports, from period three on in each regime, the frequencies with which joint plays, or "profiles of play," fall into each of the four game cells.

In the four dominant-row-strategy regimes ($\mathcal{A}v\mathcal{B}$, $\mathcal{A}v\mathcal{D}$, $\mathcal{C}v\mathcal{B}$, and $\mathcal{C}v\mathcal{D}$), a very high percentage of play occurs in the respective Nash equilibrium cells. This is not surprising: virtually every existing model of learning predicts that this outcome will be reached relatively quickly in games in which one player has a dominant strategy. But it is nevertheless reassuring: in every case, the Column player's Nash equilibrium play depends upon the Row player's type, and the players can therefore be said to have "learned" to play the NE in each of these games, as learning theories predict.

In the multiple-equilibrium regimes, $\mathcal{B}v\mathcal{B}$ and $\mathcal{D}v\mathcal{D}$, the *distinct* predictions of the Jordan model are supported by the data. Each of these regimes is a game with one mixed strategy equilibrium and two pure strategy equilibria. The regime $\mathcal{B}v\mathcal{B}$ is a symmetric coordination game in which the players have opposing preferences for the two pure strategy equilibria, as in Battle of the Sexes games. Most theories of learning do not predict which of the two pure strategy equilibria will be played in this kind of game. However, Jordan's model predicts that when the $\mathcal{B}v\mathcal{B}$ game is encountered in the environment $\mathcal{T}_1 \times \mathcal{T}_2$, with the prior beliefs induced in this experiment, the players will learn to play the profile (B,L). It is evident in Table 4 that the most frequently played profile was indeed (B,L). More striking is the fact that the other pure strategy equilibrium, (T,R), was the most *infrequently* played profile.

In the regime $\mathcal{D}v\mathcal{D}$, it is clear from Table 4 that nearly all play was at the profile (T,L), which is the unique prediction of the Jordan model. Here, however, there is more reason to expect a particular equilibrium than there is in the $\mathcal{B}v\mathcal{B}$ regime: the equilibrium (T,L) strongly Pareto dominates the other pure-strategy equilibrium, (B,R).

In the two regimes that have a unique equilibrium in mixed strategies, $\mathcal{B}v\mathcal{D}$ and $\mathcal{D}v\mathcal{B}$, all four profiles were indeed played with considerable frequency. There is no strong evidence, however, that the observed frequencies are consistent with the games' mixed strategy equilibria. Both games have the same equilibrium, in which Row plays Top with probability one-third and Column plays Left with probability one-third. Thus, mixed-strategy play should yield cell frequencies in each regime that are approximately as follows: 11% for (T,L), 22% each for (T,R) and (B,L), and 44% for (B,R). Table 5 describes goodness-of-fit tests of the hypothesis that the data were generated by this multinomial distribution, and also of the hypothesis that the data were generated by the equiprobable distribution.

As Table 5 indicates, the hypothesis that the data were generated by the mixed strategy multinomial distribution is rejected at the .10 level of significance. Furthermore, if we pool the data from the two regimes and test the hypothesis that all these data were generated by the two games' identical mixed-strategy equilibria, we reject that hypothesis at the .05 level of significance. This is because the two samples deviate from the null hypothesis in the same direction, thus increasing the value of the test statistic when pooled and, of course, the pooled sample size is larger, increasing the power of the test. Indeed, as Table 5 describes, we cannot at any reasonable significance level reject the hypothesis that the two regimes' data were generated by the same random process. For the hypothesis that the data were generated by an equiprobable random process (i.e., each profile with probability 1/4), the test statistic for each set of data (for the two regimes treated separately, and for the pooled data) is somewhat smaller than the corresponding test statistic for the mixed-strategy hypothesis; however, the equiprobable random hypothesis can be rejected for the pooled data at the .10 level of significance. Finally, we test whether the data from the two regimes come from the same distribution and find that we cannot reject this hypothesis (the p-value is 0.50).

4.2 Consistency of the Data with Path Predictions

Of course, it is quite possible that subjects "learn" to play a Nash equilibrium of the true game they are playing, but that the Jordan model is not an accurate or useful description of *how* they learn. In particular, the path of subjects' play over time may not be consistent with Jordan's model, even if the observed path does converge to a pure strategy Nash equilibrium predicted by the model. We therefore undertake an assessment of the *path consistency* of our experimental data with Jordan's model of learning.

Because of the fully rational character of Jordan's model, an assessment of the empirical accuracy of the model's path predictions is difficult. In our experimental environment, the Jordan model makes specific point predictions about behavior, and any deviation from these predictions could be construed as a rejection of the theory. This kind of rejection, however, is not very informative. Instead, we take the following, much more limited approach to evaluating the empirical accuracy of the model's period-by-period predictions: Conditional upon a player's information state – i.e., upon his type and his observations of past play by his opponent(s) – does Jordan's Bayesian model exhibit some ability to account for the observed data?

In order to answer this question we have to determine which of the choices that our experimental subjects made were in circumstances in which the Jordan model actually makes a prediction. The difficulty here is that once a player chooses an action "off the Jordan path," i.e., an action that current beliefs associate only with a type other than the player's actual type, then his co-player is very likely to eventually find himself unable to update to a well defined belief via Bayes' Rule. In our experimental environment, this theoretical inability to update typically occurs very quickly, and the model then makes no prediction about subsequent activity.

Therefore, we will consider only the first two periods of play in each regime. The inability to update cannot occur in the first two periods. In the first period, only one of the Bayesian ingredients of Jordan's model comes into play: players are predicted to play a Bayesian Nash equilibrium, but they have not yet observed any play and so have no information that they can use to update their beliefs. And at the second period, beliefs updated on the basis of first period observation are always well defined, because each of the two actions available to each player is part of a first-period Bayesian Nash equilibrium for at least one of his types. Further, in accordance with the myopic-play assumption of the Jordan model, we assume that a player who has made a "wrong" move at the first period is subsequently predicted to make the "correct" move for his information state at the second period. (We are effectively assuming that, because of the experiment's rematching of subjects at period two, a player will expect that his co-player in period two has observed the "correct" play at period one.) Finally, we will consider only the choices by subjects who did not have a dominant strategy (thus ruling out Row types $\mathcal{A}$ and $\mathcal{C}$); recall that the subjects who had a dominant strategy chose that strategy 99% of the time.

Table 6 presents a summary of all observed play in the first period by players who did not have a dominant strategy (all Column players, and all type $\mathcal{B}$ and $\mathcal{D}$ Row players). The Jordan model predicts that all play will be at the Bayesian Nash equilibrium, which corresponds to the cells of Table 6 containing the bold-faced entries. We ask whether the data presented in Table 6 are likely to have been generated by an equiprobable binomial process (i.e., a process that, for each player, yields each of his strategies with probability 1/2). The appropriate chi-square test rejects the hypothesis of equiprobable play at levels of significance of 1% for Row players and 10% for Column players, as indicated in Table 6. Alternatively, consider the hypothesis that Row players of a given type table play Top and Bottom according to a binomial process, and that the two binomial processes (the one for $\mathcal{B}$ players and the one for $\mathcal{D}$ players) are the same – i.e., that they have the same binomial probability. This hypothesis is rejected at significance level 5%, and the parallel hypothesis for the Column players is rejected at significance level 10%. Thus, we conclude that Bayesian Nash equilibrium does have some explanatory power for period 1 in our experimental environment.

In the second period the Jordan model assumes that the players will combine Bayesian updating of beliefs with Bayesian Nash equilibrium play. Table 7 presents a summary of observed second period play by those subjects whose payoff tables did not have a dominant strategy. The entries in the left half of Table 7 indicate, for Column subjects of a given type ($\mathcal{B}$ or $\mathcal{D}$) whose first period opponents played a given strategy (T or B), how many of those Column subjects played Left and how many played Right in the second period. The analogous figures for the Row subjects are the entries in the right half of Table 7. (Note, however, that the Row data is too sparse to draw any conclusions.) The Jordan model predicts that all play will be in the cells of Table 7 that contain the bold-faced entries. Note that for all four Column information states, and for the Type $\mathcal{D}$ Row subjects, the bold-faced entries are all substantially larger than the corresponding unbolded entries. For each of the Column types, $\mathcal{B}$ and $\mathcal{D}$, Chi-square tests reject the hypothesis that these data were generated by an equiprobable process (i.e., Left and Right equally probable in each of the two information states, T and B). The parallel test when we pool the $\mathcal{B}$ and $\mathcal{D}$ types rejects at a significance level of 2%. Thus, Table 7 yields informally the same conclusion as Table 6 – there is some limited support for Jordan's Bayesian model in the observed path of play in the first two periods.

4.3 Alternative Learning Models

It is important to ask not only whether a model is consistent with observation — does it predict or explain well? — but also whether it performs better in this respect than alternative models. Although we did not design our experiment with this kind of comparison in mind, we address the issue briefly, by asking whether the experimental data provide more support for Jordan's Bayesian model than for the Best Response model or the Fictitious Play model of learning in games.

The comparison of the models that we would like to make would be a period-by-period comparison of the experimental data with the path predictions of each of the alternative models. For several reasons, however, the data from our experiment provide only an extremely limited comparison. For example, the Best Response and Fictitious Play models make no prediction for period one, so we have nothing to compare with the discussion of Table 6 that we provided above for the Jordan model. For subsequent periods, the models all make specific predictions (predictions conditional on first-period play in the case of Fictitious Play and Best Response), but the data have further limitations, beyond even the difficulties described in the preceding section. For the games $\mathcal{B}$ *vs* $\mathcal{D}$ and $\mathcal{D}$ *vs* $\mathcal{B}$, as we have already indicated, there is simply not enough data to draw any conclusions. For all of the remaining games, the Jordan model predicts that play in period two will be repeated at all subsequent periods. However, in the games with a dominant Row stategy, the predictions of the Best Response and Fictitious Play models in all periods after period one duplicate those of the Jordan model exactly. Thus, there is nothing to choose between the models in these games.

This leaves just the games $\mathcal{B}$ *vs* $\mathcal{B}$ and $\mathcal{D}$ *vs* $\mathcal{D}$, which are the games with multiple equilibria. Here, as we have already seen, the Jordan model predicts which equilibrium will be eventually played, as well as the path by which play will evolve over time, and there is some agreement of the experimental observations with these predictions. Without any information about first-period play, the Best Response and Fictitious Play models make no prediction about the equilibrium that will eventually be reached, so in this regard the Bayesian model makes sharper predictions, predictions that are consistent with the data. On the other hand, once play at the first period is known, the Best Response and Fictitious Play models also make precise predictions about subsequent play in this experiment's true games. Moreover, these two models continue to make well defined predictions even after players (subjects) have played in ways that are inconsistent with previous predictions, in contrast with the Bayesian model. As Table 7 indicates, the first-period play that actually occurred in the experiment produced second-period predictions by the Best Response and Fictitious Play models that were in almost all cases exactly the same as the Jordan model's prediction. Again, then, there is little scope for saying that any one of the models is better than the others at accounting for the data.[1]

5. Conclusions

In order to conduct a test of the empirical implications of Jordan's model of Bayesian learning in games, and to compare the model to alternative models of learning, the experimenter must construct an incomplete information environment that satisfies several desiderata. Chief among our desiderata in developing the experiment we have reported here were (1) minimization of the number of both multiple and mixed-strategy Bayesian Nash equilibria, especially in the early periods; (2) minimization of the number of games with multiple or mixed-strategy Nash equilibria of the true game; and (3) inclusion of sufficiently many plays in each fixed "regime" to be able to assess whether convergence to the Nash equilibrium predicted by the Jordan model had occurred.

The experiment we have described provides support for the idea that equilibrium play can be attained through repeated play in an environment of incomplete information — *i.e.*, the experiment provides support for the "learning approach" to rationalizing equilibrium analysis. In particular, the cell frequencies in Table 4 support the conclusion that play can converge to a pure strategy Nash equilibrium without the players having common knowledge of one another's payoffs, beliefs, and rationality.

The experiment was also designed to test Jordan's specific model of the path by which equilibrium play is learned. Our experiment provides very limited support for Jordan's model, but is largely unable to distinguish between the Jordan model's predictions and those of such other learning models as best repsonse and fictitious play. In the games with a unique equilibrium in pure strategies, all three models' predictions about eventual play were supported. The results from the $\mathcal{B}v\mathcal{B}$ games, which have three Nash equilibria (two of which are in pure strategies), provide a very suggestive and sharp instance of empirically successful equilibrium selection by the Jordan model but not by the other two models. The data that appear in Table 4 provide support for the Jordan model's prediction that play will be at the equilibrium (B,L) and not at the other pure strategy Nash equilibrium, (T,R). The data therefore constitute an empirical demonstration that players' prior beliefs, combined with Bayesian learning from observed play, can in some cases provide a solution to the equilibrium selection problem.

FOOTNOTE

1. Cox, Shachat, and Walker (1995) describes an experiment that was designed to maximize the extent to which alternative learning models make different predictions about play.

REFERENCES

Cox, James C., Jason Shachat, and Mark Walker (1995): "An Experimental Test of Bayesian vs. Adaptive Learning." Photocopy.

Jordan, James S. (1991): "Bayesian Learning in Normal Form Games," *Games and Economic Behavior*, 3, 60-81.

Kalai, Ehud, and Ehud Lehrer (1993): " Rational Learning Leads to Nash Equilbrium," *Econometrica*, 63, 1019-1045.

Table 1: Equilibria and Predicted Path for Each Game

True Game	Nash Equilibria	Path of Play Predicted by Jordan Model t=1	t=2	t=3,4,5,...
A vs ***B***	(T,R)	(T,R)	(T,R)	(T,R) ...
A vs ***D***	(T,L)	(T,L)	(T,L)	(T,L) ...
C vs ***B***	(B,L)	(B,R)	(B,L)	(B,L) ...
C vs ***D***	(B,R)	(B,L)	(B,R)	(B,R) ...
B vs ***B***	(B,L), (T,R), (1/3, 1/3)*	(B,R)	(B,L)	(B,L) ...
D vs ***D***	(T,L), (B,R), (1/3, 1/3)*	(T,L)	(T,L)	(T,L) ...
B vs ***D***	(1/3, 1/3)*	(B,L)	(T,R)	(1/3, 1/3)* ...
D vs ***B***	(1/3, 1/3)*	(T,R)	(B,R)	(1/3, 1/3)* ...

* A fraction denotes a mixed-strategy equilibrium: Row's fraction is the probability with which he plays Top; Column's fraction is the probability with which he plays Left.

Table 2: Experimental Regimes

(Row player type vs. Column player type)

Session 1:	***C*** vs ***B***	***A*** vs ***B***	***D*** vs ***D***	***A*** vs ***D***	
Session 2	***A*** vs ***B***	***D*** vs ***D***	***B*** vs ***D***	***C*** vs ***B***	***D*** vs ***D***
Session 3	***A*** vs ***D***	***C*** vs ***B***	***C*** vs ***B***	***A*** vs ***B***	***A*** vs ***B***
Session 4	***B*** vs ***B***	***D*** vs ***B***	***C*** vs ***D***	***D*** vs ***D***	***B*** vs ***B***

Table 3: Salient Payoffs

	Low Payoff	Mean Payoff	High Payoff
Session 1:	$17.67	$21.99	$26.00
Session 2:	$15.90	$16.96	$18.20
Session 3:	$12.70	$15.73	$19.10
Session 4:	$11.10	$12.95	$13.90

Table 4: Cell Frequencies for Periods Three and Later

(Bold face indicates a unique prediction by the Jordan model)

Cell	***A*** vs ***B***	***A*** vs ***D***	***C*** vs ***B***	***C*** vs ***D***	***B*** vs ***B***	***D*** vs ***D***	***B*** vs ***D***	***D*** vs ***B***
(T,L)	3%	**91%**	0%	0%	20%	**88%**	14%	18%
(T,R)	**96%**	4%	0%	0%	12%	5%	23%	28%
(B,L)	0%	4%	**96%**	17%	**40%**	5%	32%	22%
(B,R)	1%	0%	4%	**83%**	28%	2%	31%	32%
# of Obs.	312	156	312	78	156	312	78	78

Table 5: Hypothesis Tests for Equilibrium Play in Regimes with a Unique Mixed Strategy Equilibrium

Game	Null	Alternative	Test Statistic	Distribution	P-value
B vs ***D***	Cell frequencies are generated by multinomial distribution implied by mixed strategy N.E.	Cell frequencies are not generated by multinomial distribution implied by mixed strategy N.E.	7.61	Chi-square with 3 d.o.f.	0.05
B vs ***D***	Cell frequencies are generated by a uniform multinomial distributon	Cell frequencies are not generated by a uniform multinomial distributon	6.41	Chi-square with 3 d.o.f.	0.09
D vs ***B***	Cell frequencies are generated by multinomial distribution implied by mixed strategy N.E.	Cell frequencies are not generated by multinomial distribution implied by mixed strategy N.E.	6.77	Chi-square with 3 d.o.f.	0.08
D vs ***B***	Cell frequencies are generated by a uniform multinomial distributon	Cell frequencies are not generated by a uniform multinomial distributon	5.90	Chi-square with 3 d.o.f.	0.12
B vs ***D*** & ***D*** vs ***B*** (pooled)	Cell frequencies are generated by multinomial distribution implied by mixed strategy N.E.	Cell frequencies are not generated by multinomial distribution implied by mixed strategy N.E.	11.73	Chi-square with 3 d.o.f.	0.01
B vs ***D*** & ***D*** vs ***B*** (pooled)	Cell frequencies are generated by a uniform multinomial distributon	Cell frequencies are not generated by a uniform multinomial distributon	7.85	Chi-square with 3 d.o.f.	0.05
B vs ***D*** & ***D*** vs ***B*** (pooled)	Cell frequencies in both games are generated by the same distribution	Cell frequencies in both games are generated by different distributions	2.36	Chi-square with 3 d.o.f.	0.50

Table 6: Period 1 Actions and Hypothesis Tests

(The predictions by the Jordan model are in Bold Face)

Row Subjects

	Type ***B***	Type ***D***
Top	8	**23**
Bottom	**10**	7
Total	18	30

Test of the hypothesis that subjects' choices of Top and Bottom are equiprobable:

Chi-square Statistic:	8.7556
degrees of freedom:	1
P-value:	0.0031

Test of the hypothesis that subjects choose Top with the same probability whether they are Type ***B*** or Type ***D:***

Z-statistic:	-2.26
P-value:	0.0119

Column Subjects

	Type ***B***	Type ***D***
Left	27	**27**
Right	**39**	21
Total	66	48

Test of the hypothesis that subjects' choices of Left and Right are equiprobable:

Chi-square Statistic:	2.9318
degrees of freedom:	1
P-value:	0.0868

Test of the hypothesis that subjects choose Left with the same probability whether they are Type ***B*** or Type ***D:***

Z-statistic:	-1.62
P-value:	0.0527

Table 7: Period 2 Actions and Hypothesis Tests

The Number of Subjects in Each Information State who Played Each Action
(An information state consists of the subject's type and his opponent's period-one action)

Actions predicted by the Jordan model are in Bold face

Row Subjects

	(***B*** ;L)	(***B*** ;R)	(***D*** ;L)	(***D*** ;R)
Top	**3**	6	**13**	6
Bottom	6	**3**	1	**10**
Total	9	9	14	16

Column Subjects

	(***B*** ;B)	(***B*** ;T)	(***D*** ;B)	(***D*** ;T)
Left	**18**	11	4	**29**
Right	13	**24**	**11**	4
Total	31	35	15	33

Test of the hypothesis that Column
subjects' choices were equiprobable
in each of the four information states:

Chi-square Statistic:	5.64
degrees of freedom:	1.00
P-value:	0.02

Test of the hypothesis that Column
subjects' choices were equiprobable
in the two Type ***B*** information states:

Chi-square Statistic:	32.82
degrees of freedom:	1.00
P-value:	0.00

Test of the hypothesis that Column
subjects' choices were equiprobable
in the two Type ***D*** information states:

Chi-square Statistic:	38.46
degrees of freedom:	3.00
P-value:	0.00

Figure 1: Normalized Payoff Tables

	L	R
T	a	b
B	0	0

Row Player's
Payoff Function π_1

	L	R
T	α	0
B	β	0

Column Player's
Payoff Function π_2

	L	R
T	a,α	$b,0$
B	$0,\beta$	$0,0$

The Game
π_1 vs π_2

Figure 2

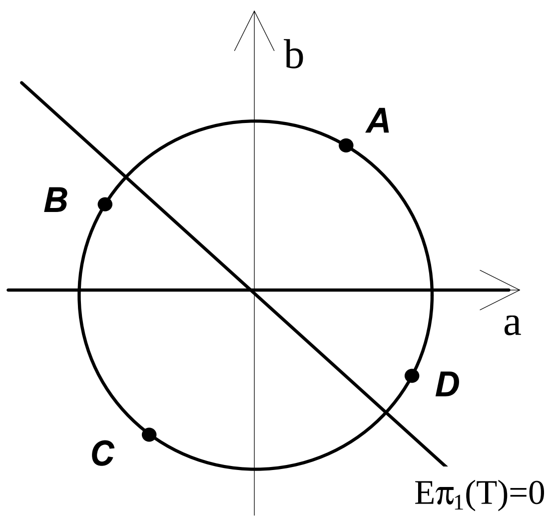

Figure 2a

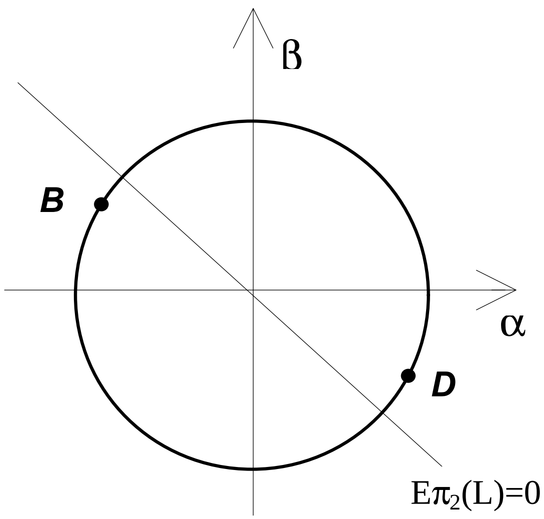

Figure 2b

Figure 3: Conversion to Nonnegative Payoff Tables

Type B Row players, for example, have this payoff function:

	L	R
T	0	3
B	2	2

... instead of this one:

	L	R
T	-2	1
B	0	0

Figure 4: Induced Prior Beliefs

(Types and Probabilities)

The Row Player

Type A

1	2
0	0

Probability: 3/8

Type B

0	3
2	2

Probability: 1/8

Type C

1	0
2	2

Probability: 3/8

Type D

3	0
1	1

Probability: 1/8

The Column Player

Type B

0	2
3	2

Probability: 1/2

Type D

3	1
0	1

Probability: 1/2