GAMES AND ECONOMIC BEHAVIOR 5, 205–222 (1993)

Fictitious Play: A Statistical Study of Multiple Economic Experiments

RICHARD T. BOYLAN*

Olin School of Business, Washington University, St. Louis, Missouri 63130

AND

MAHMOUD A. EL-GAMAL*

Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125

Received June 17, 1991

We propose a methodology for combining a number of experiments to make an overall assessment of the empirical justifiability of certain economic hypotheses. We illustrate our methodology by an application to models of learning in repeated games. The application of our proposed (purely Bayesian) procedure allows us to combine two sets of experiments (for a total of nine experiments) to update our beliefs on the relative justifiability of the Cournot and the fictitious play learning hypotheses. The experiments we analyzed indicate that the fictitious play hypothesis is impressively more likely than the Cournot hypothesis. *Journal of Economic Literature* Classification Numbers: 026,211,215.

1. INTRODUCTION

In recent years, there has been a marked growth in the number of experimental studies in economics, and their influence on the field can no longer be denied.[1] The interaction between experimental studies and other parts of the discipline has also been flourishing, but we believe that it has not yet reached its potential. We believe that the main reason for the

* We thank Bob Forsythe and Gary Miller for giving us access to their data. We thank an anonymous referee and the editor in charge for comments that helped substantially to improve the paper.

[1] A review of experimental findings can be found in Hoffman and Spitzer (1985), Plott (1982), and Smith (1982).

0899-8256/93 $5.00

slow development of those interactions is our inability to integrate a large number of experimental studies systematically (in some rigorous statistical manner) to develop stylized facts that can inspire and benefit from work in other subfields of economics. It is true that a number of stylized facts have already been established by experimental studies and have been taken into account by modeling economists, but the progress of the science can be greatly enhanced by a systematic procedure of generating such stylized facts. This is the issue that we address in this paper. It was clear to us from the start that no classical procedure can be used to combine the results of different experiments in generating stylized facts. That is the case since the different experiments have completely different designs and were performed to investigate different issues. To complicate things further, one cannot ignore the fact that experimenters are aware of the other experiments that their colleagues have run, and hence a classical procedure will need to take into consideration what amounts to pretesting of the hypotheses in question, a task that defies even the most capable of statisticians.

We analyze the data in a Bayesian fashion, endowing ourselves with a belief on the relative validity of a number of possible theories of human economic behavior and updating our beliefs using the experimental results available to us. Merits and demerits of our Bayesian procedure as opposed to some classical procedures will be discussed in a later section.

We believe that the best way to illustrate a new procedure is by example, and for this paper, we chose to analyze different learning procedures in repeated games. The motivation for the learning literature is to explain why individuals should be expected to select strategies that correspond to Nash equilibria. The answer suggested by that literature is that the game is played repeatedly. At each stage, individuals select the action that maximizes their expected payoff in that round. Players start with beliefs on their opponents' possible actions, and update their beliefs as they play each stage of the game and observe their opponents' actual actions. A Nash equilibrium is defined by a set of self-fulfilling beliefs.

In this paper, we consider two of the most popular learning hypotheses in repeated games. The first hypothesis was proposed by Cournot and proposes that each agent assumes that her opponents will choose the same action they chose in the previous period (see Moulin, 1985). The second hypothesis that we analyze is the so-called fictitious play hypothesis (see Brown, 1951; Robinson, 1951; Shapley, 1962; Brock *et al.*, 1988). The fictitious play hypothesis proposes that an individual has Dirichlet priors over her opponents' strategies and that at each round the player updates her beliefs according to Bayes' rules. It turns out that these assumptions are equivalent to requiring that each player selects the strategy that maximizes her payoff given beliefs that correspond to a convex combination of the initial beliefs and the empirical distribution. The experimental literature investigating different learning algorithms is very sparse, and limited to an experiment-by-experiment study. We illustrate our proposed methodology by starting with prior probabilities about the relative plausibility of the two learning hypotheses described above. We use a simulation technique to compute the likelihood functions (the probability of observing the data in each experiment under each of the two hypotheses), and then use those likelihood functions to do Bayesian updating of our priors for each individual experiment, and then for all nine experiments combined.

Figure 1 summarizes the results of this paper. When we start the analysis, we need a prior on the relative likelihoods of the Cournot and the fictitious play hypotheses. A prior $p \in [0, 1]$ is the probability we assign to the fictitious play hypothesis as opposed to the Cournot hypothesis before we observe any data. After we observe the data, we use the relative likelihoods of the two hypotheses (the ratio of the probabilities of observing the realized data given each of the two hypotheses) and the prior to obtain the posterior probability that the fictitious play hypothesis holds (another number between 0 and 1). In our analysis, we introduce a smearing parameter $\varepsilon \in [0, 1]$ representing the probability for each person at each move of each game to have a tremble and make a completely random move (uniformly over available actions). The need for a smearing parameter and the particular configuration we choose are discussed in detail in later sections. Figure 1 shows a plot of our posterior probability (using all nine experiments) on the fictitious play hypothesis as a function of possible priors and values of the smearing parameter. It is clear that for all priors, and for all but very high smearing parameters (at $\varepsilon = 1$, the data has no influence on the posterior), the smeared fictitious play hypothesis is impressively more likely than the smeared Cournot hypothesis.

The rest of this paper builds up and justifies the necessary machinery to achieve Fig. 1. Section 2 describes the experiments that we analyze.

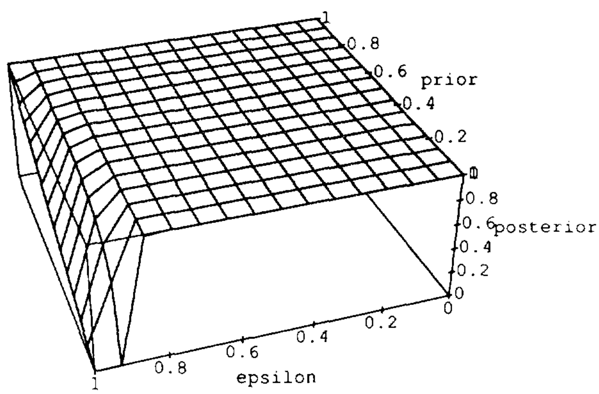


FIG. 1. Posterior probability of agents playing according to fictitious play using all nine experiments.

Section 3 describes and justifies the actual statistical and numerical techniques that we follow. Section 4 discusses the advantages and disadvantages of our methodology, and Section 5 contains some concluding remarks. The paper ends with a series of nine appendixes each describing an experiment and presenting our statistical analysis of the data it generated.

2. Description of the Experiments and the Appendices

The first set of experiments that we analyze were run by Knott and Miller (1987). In this series of three experiments (labeled A, B, and C), individuals were matched in pairs and played the games reproduced in Figs. 2, 7, and 12, respectively, 10 times. The second set was run by Cooper *et al.* (1990). In each experiment of that second set, there are 11 players. Each player plays twice against each of the other players, where the matching in each round is determined at random. Agents do not know the identity of the player they are matched with and after each play they find out which strategy the opponent selected. We analyze 6 of the experiments run by the authors. These experiments are labeled Experiments 3 through 8, and the payoff matrices are depicted in Figs. 17, 22, 27, 32, 37, and 42.

There are nine appendixes corresponding to the nine experiments that we analyze. The appendixes are labeled A, B, C (for Knott and Miller's experiments A, B, and C, respectively), and 3 through 8 (for Cooper *et al.*'s experiments 3–8, respectively). Each appendix contains five figures. First, the normal form matrix of the game played in the corresponding experiment is displayed together with notes describing the pure strategy Nash equilibria of the game, and when appropriate, which of them we expect to be the limit strategy for most players. The statistical analysis of the individual experiment shown below is then discussed. The next two figures show the simulated frequencies with which players would select the various actions in each period of the game if they used Cournot and fictitious play updating, respectively. To obtain these simulated frequencies we simulated 1,000 ensembles of players (2,000 individuals in experiments A–C and 11,000 individuals in experiments 3–8) and endowed them with randomly drawn (uniform over the strategy simplex) initial priors over their opponents' possible actions.[2] We then let each pair go through the identical design that was implemented in the experiment (10 stages of the game with the same opponent for experiments A–C, and the exact matching scheme realized in experiments 3–8) and where they were made to update according to the Cournot and fictitious play rules, respectively.

[2] We have checked the sensitivity of those simulated frequencies to the size of our Monte Carlo simulations, and found that 10,000 ensembles led to almost identical frequencies.

The fourth figure shows the actual observed frequencies of play of each action in each period. The fifth figure shows our posterior belief on fictitious play against Cournot updating using the likelihood function for that experiment alone. The posterior is plotted as a function of all possible priors and the level of smearing ε (to be explained in the next section). Figure 1 is simply the combination, through Bayesian updating, of the fourth figures of all nine appendixes to produce the overall posterior belief.

3. Econometric Analysis

In the previous section, we described how we obtained the simulated frequencies of play of each of the available actions in each of the time periods under the two competing hypotheses. Let us now introduce some notation. Let $p_c^t = 1 - p_f^t$ be the experimenter's subjective probability at time t that individuals act according to the Cournot process. Also, let $q_c^t = 1 - q_f^t$ be the probability of the observed strategy choices at time t given that the individuals act according to the Cournot process. The experimenter updates her beliefs according to Bayes' rule:[3]

$$p_c^t = \frac{p_c^0 q_c^1 q_c^2 \cdots q_c^t}{p_c^0 q_c^1 q_c^2 \cdots q_c^t + p_f^0 q_f^1 q_f^2 \cdots q_f^t},$$

$$p_f^t = \frac{p_f^0 q_f^1 q_f^2 \cdots q_f^t}{p_c^0 q_c^1 q_c^2 \cdots q_c^t + p_f^0 q_f^1 q_f^2 \cdots q_f^t}.$$

Now, when we compute q_i^t for $i = $ f, c, if we use the simulated frequencies depicted in the appendices, we run into the *zero likelihood problem*. This problem is common in experimental economics and occurs when some of the observed data have a zero probability of occurring under all the hypotheses that we consider. For instance, in the Cooper *et al.* experiments there are observations of agents playing strictly dominated actions which cannot be justified under any beliefs. The zero likelihood problem has been discussed at length in El-Gamal *et al.* (1992), and a number of solutions that exist in the literature were discussed. The solution we use here is similar in spirit to those used in McKelvey and Palfrey (1992) and El-Gamal *et al.* (1992). We include in our models a small probability of a tremble (a completely random action) taking place. The two smeared models we compare are then

$$\text{Model 1: } a' = \begin{cases} a_c & \text{with probability } 1 - \varepsilon \\ a & \text{with probability } \dfrac{\varepsilon}{A}, \end{cases}$$

[3] It trivially follows that the order with which the experiments are analyzed does not affect the belief of the experimenter.

$$\text{Model 2: } a' = \begin{cases} a_f & \text{with probability } 1 - \varepsilon \\ a & \text{with probability } \dfrac{\varepsilon}{A}, \end{cases}$$

where a' is the action chosen by the experimental subjects, a_c is the action that maximizes their expected payoff given their beliefs that are updated according to the Cournot procedure, a_f is the same as a_c with fictitious play updating, and $a \in \{1, 2, \ldots, A\}$ is any strategy.

In computing the q_i^t's, we treat the actual action of an agent in a period as a random draw from the probability distribution obtained from the simulated experiments with probability $(1 - \varepsilon)$ and as a uniform random draw with probability ε. Let the observed data points be indexed by $(n, t) \in \{1, 2, \ldots, N\} \times \{1, 2, \ldots, T\}$, and let the simulated probability under the pure version of Model i of action a in period t be $q_{i,a}^t$. Then we compute the q_i^t needed for our Bayesian updating as

$$q_i^t = \prod_{n=1}^{N} \prod_{t=1}^{T} \left((1 - \varepsilon)q_{i,a_{n,t}}^t + \varepsilon \frac{1}{A}\right),$$

where $a_{n,t}$ is the actual action chosen by agent n in period t. Note that our smearing procedure guarantees that q_i^t is always positive, and hence avoids the zero-likelihood problem discussed above.

4. Remarks on Our Econometric Procedure

Since we are advocating the use of our new methodology for the analysis of a number of economic experiments, we should point out its main advantages and warn against its main disadvantages. This section will deal with those main advantages and disadvantages that are inherent in the general statistical methodology, as well as those that are specific to our implementation in this paper.

4.1. *Advantages of Our Procedure*

- As we have argued in the introduction, classical methods cannot be used to combine such a heterogeneous collection of experiments whereas the Bayesian procedure we used can. This is a general property of Bayesian methods: the order in which evidence appears and the source of the evidence are irrelevant.
- The type of horse race we ran between our two hypotheses could just as easily be run with any number of hypotheses. The relative likelihood of each of the hypotheses is still the likelihood of each of them

divided by the sum of the likelihoods. As the relative belief (prior times the relative likelihood) for some hypothesis becomes negligible, it can be dropped out of the race without altering the methodology. Similarly, new hypotheses that suddenly gain credibility can be added to the race.

• In addition to the previous point, the nature of this Bayesian procedure also keeps all hypotheses alive and can be resurrected at all times. This is in contrast to classical hypothesis testing procedures which choose one model over all others, and in which dead models cannot be resurrected in a probabilistically consistent manner.

• The computation of the likelihoods of all possible observations under each of the competing hypotheses can clearly be done prior to running the experiments. Hence, if we were to run more experiments à la Cooper *et al.* to distinguish between our two hypotheses in this paper, we would choose the design of their experiment 6, where there is a very strong separation between the two experiments, and avoid the design of experiment 7, where there is practically none.

4.2. *Shortcomings of Our Procedure*

• The choice to draw initial beliefs uniformly for our simulations seems rather arbitrary. It can be justified on the basis of being the least informative prior (the one which maximizes entropy) on the players' beliefs. It is, however, conceivable that different games will give rise to different initial distributions of beliefs.

• In our Bayesian approach, the parameter ε is what is commonly called a *nuisance parameter*, and so we chose to be agnostic as to its value. Two other alternatives are available. One would be to follow a classical procedure and label ε an irrationality parameter as in McKelvey and Palfrey (1992) and find its estimate under the maintained hypothesis of one of our models. The alternative is to follow a Bayesian methodology as in El-Gamal *et al.* (1992) and start with priors on the nuisance parameter ε, integrate it out to compute the likelihoods, and then update our prior about it under each of the competing models. In this paper, we decided that by displaying the posterior beliefs on the competing hypotheses at all values of our prior beliefs and the parameter ε, we allow the readers to integrate with respect to whatever prior beliefs they may choose.

• The choice to maintain a constant ε can be criticized on the basis that agents should be learning over time. Given the shortness of the time series component of the data we use, however, we cannot extract any parametric form for such a learning-by-doing curve, and nonparametrics with such small data samples would be extremely unreliable.

• It is very difficult to decide on a comprehensive set of hypotheses that will please all readers of a survey. But since our Bayesian procedure allows us to add hypotheses as wished, this problem is one that should concern only sociologists of science.

• The set of available experiments may not be well suited for distinguishing between the hypotheses that interest us since it was run for a different purpose. This suggests that we cannot stop running experiments and concentrate on analyzing the very large number of experiments that have already been run. The need for well designed experiments to update our beliefs about various hypotheses continues to be very strong.

• There is a well known problem that arises in the context of other methods of combining statistical results (e.g., El-Gamal, 1992); meta-analsis (see Wolf, 1986); and the consensus literature (see Genest and Zidek 1986). For instance, there is always a bias towards significant and surprising findings in published research, which will influence the direction of the overall analysis. This influence can be quite harmful if the "significance" of the results for the experimenter means that it favors one particular set of hypotheses.

• Well designed and run experiments are given the same weight in this methodology as badly designed and/or run experiments. This problem was dealt with in El-Gamal (1992) by allowing different weights to be assigned to different studies, but the decision on those weights can be a very difficult problem.

5. Concluding Remarks

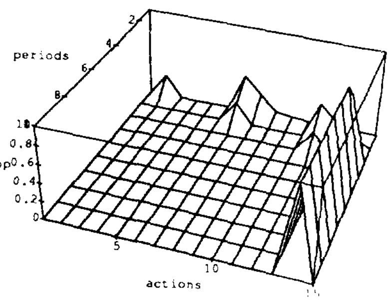
In this paper, we have proposed and illustrated a fully Bayesian procedure for updating our beliefs on a number of economic hypotheses. Of the nine experiments that we analyzed in the appendixes, some unequivocally favor the Cournot learning hypothesis and some unequivocally favor the fictitious play learning hypothesis. No classical statistical procedure will allow us to analyze all nine experiments to decide on what we have learned from all of them combined. This is the case due to the different designs and purposes of the experiments in question. Bayesian procedures, on the other hand, require only the computation of likelihood ratios for the models in question for each of the experiments. Our overall Bayesian analysis of the nine experiments shows that starting from any prior on the relative validity of the two learning hypotheses, we end up believing that the fictitious play learning hypothesis is infinitely more likely. The stylized facts derived with this approach will hopefully generate more productive interaction between theory and experimentation. We must also warn, however, that one must be very careful when applying a general technique like ours not to fall into too many of the problems outlined in the previous section lest one's beliefs be wrongly biased in favor of some hypotheses.

Appendix A. Experiment A of Knott and Miller[4]

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
S1	100	100	75	85	85	90	90	90	90	90	90	90	90	90	90
S2	200	120	100	100	100	100	100	100	100	100	100	100	100	100	100
S3	0	140	125	115	115	110	110	110	110	110	110	110	110	110	110
S4	-100	160	150	130	125	120	120	120	120	120	120	120	120	120	120
S5	-200	180	175	145	135	130	130	130	130	130	130	130	130	130	130
S6	-300	200	200	160	145	140	140	140	140	140	140	140	140	140	140
S7	-400	220	220	175	155	155	150	150	150	150	150	150	150	150	150
S8	-450	0	250	190	165	160	160	160	160	160	160	160	160	160	160
S9	-475	-100	0	205	175	170	170	170	170	170	170	170	170	170	170
S10	-500	-125	-100	0	185	180	175	175	180	180	180	180	180	180	180
S11	-525	-150	-125	-100	0	190	180	180	190	190	190	190	190	190	190
S12	-550	-175	-150	-125	-100	0	195	190	195	195	200	200	200	200	200
S13	-575	-200	-175	-150	-125	-100	0	250	200	200	210	210	210	210	210
S14	-600	-225	-200	-175	-150	-125	-100	0	250	250	250	250	250	220	220
S15	-625	-250	-225	-200	-175	-150	-125	-100	80	100	120	140	160	180	200

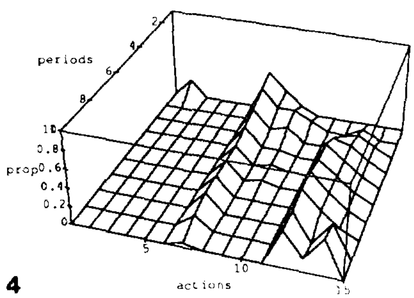
FIG. 2. Payoff matrix for Knott and Miller's experiment A.

- (S14,S14) is the unique pure strategy equilibrium.
- It is clear that both Cournot updating and fictitious play will eventually lead all the players to play the pure strategy Nash equilibrium. Indeed, when we ran fictitious play for 100 periods, all the mass converged to (S14,S14). Comparing the simulated frequencies from the two models with 10 periods of data, it is clear that the fictitious play hypothesis leads to much slower convergence to the pure strategy Nash equilibrium, and hence mimics the noisy data much better than the Cournot hypothesis.



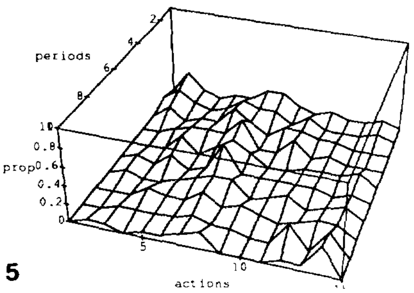
3

FIG. 3. Cournot simulated frequencies.



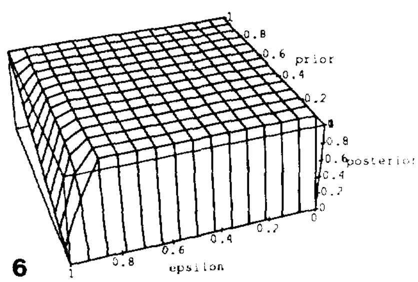
4

FIG. 4. Fict. play simulated frequencies.



5

FIG. 5. Observed frequencies.



6

FIG. 6. Posterior on fictitious play.

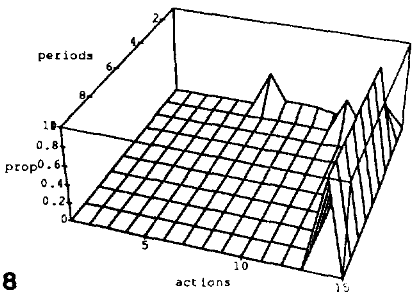
[4] In Appendixes A–C, the matrices give the payoff of the row player, while the transpose of the matrices give the payoff of the column player.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
S1	100	100	90	90	90	90	80	-100	-125	-150	-175	-200	-225	-250	-625
S2	800	120	100	100	100	100	100	0	-100	-125	-150	-175	-200	-225	-600
S3	700	400	125	110	110	110	110	110	0	-100	-125	-150	-175	-200	-575
S4	600	450	350	130	120	120	120	120	120	0	-100	-125	-150	-175	-550
S5	500	500	375	150	135	130	130	130	130	130	0	-100	-125	-150	-525
S6	400	600	400	200	200	140	140	140	140	140	50	0	-100	-125	-500
S7	350	800	425	210	210	150	150	150	150	150	100	50	0	-100	-475
S8	200	300	450	250	220	160	160	160	160	160	160	120	50	0	-450
S9	100	200	240	400	230	170	170	170	170	170	170	140	100	20	-400
S10	90	100	200	200	300	200	180	180	180	180	180	160	150	40	-300
S11	65	65	65	100	100	280	200	190	190	190	190	180	175	60	-200
S12	60	60	60	60	80	100	250	200	200	200	200	200	200	80	-100
S13	55	55	55	55	55	55	210	220	210	210	210	210	210	100	0
S14	50	50	50	50	50	50	50	50	220	220	220	220	220	220	220
S15	40	40	40	40	40	40	40	60	80	100	120	140	160	180	200

FIG. 7. Payoff matrix for Knott and Miller's experiment B.

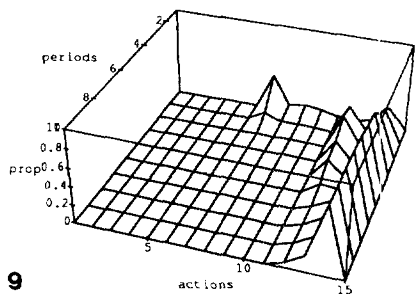
• (S14,S14) is the unique pure strategy equilibrium.

• It is clear that both Cournot updating and fictitious play will eventually lead all the players to play the pure strategy Nash equilibrium. Indeed, when we ran fictitious play for 100 periods, all the mass converged to (S14,S14). Comparing the simulated frequencies from the two models with 10 periods of data, it is clear that the fictitious play hypothesis leads to much slower convergence to the pure strategy Nash equilibrium, and hence mimics the noisy data much better than the Cournot hypothesis for low levels of ε.



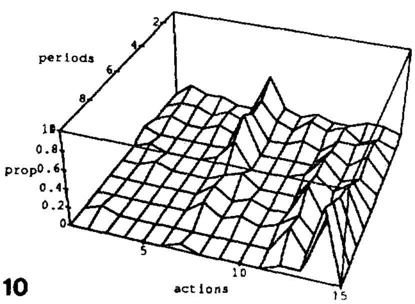
8

FIG. 8. Cournot simulated frequencies.



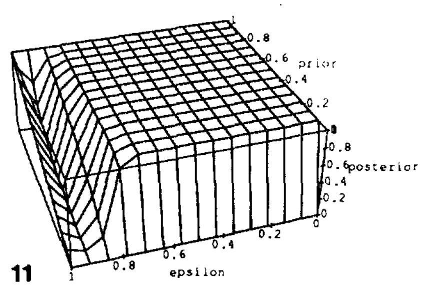
9

FIG. 9. Fict. play simulated frequencies.



10

FIG. 10. Observed frequencies.



11

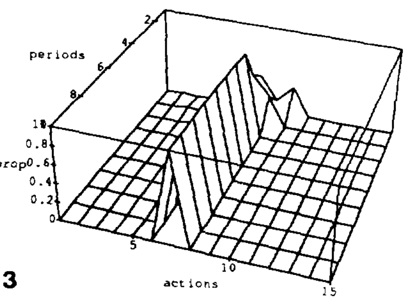
FIG. 11. Posterior on fictitious play.

APPENDIX C. EXPERIMENT C OF KNOTT AND MILLER

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
S1	100	75	85	85	90	90	90	90	90	90	90	90	90	90	90
S2	350	120	100	100	100	100	100	100	100	100	100	100	100	100	100
S3	300	140	125	115	115	110	110	110	110	110	110	110	110	110	110
S4	275	130	150	130	125	120	120	120	120	120	120	120	120	120	120
S5	250	120	125	145	135	130	130	130	130	130	130	130	130	130	130
S6	200	110	110	125	145	140	140	140	140	140	140	140	140	140	140
S7	100	100	100	100	100	155	150	300	350	250	150	150	150	150	150
S8	-350	0	50	50	75	100	125	250	400	450	450	160	160	160	160
S9	-400	-100	0	25	50	75	100	200	300	425	400	375	350	325	170
S10	-500	-125	-100	0	25	50	75	175	225	325	375	350	325	300	300
S11	-525	-150	-125	-100	0	25	50	125	200	225	350	325	300	275	250
S12	-550	-175	-150	-125	-100	0	25	100	175	200	200	250	250	250	225
S13	-575	-200	-175	-150	-125	-100	0	75	150	175	175	175	200	200	200
S14	-600	-225	-200	-175	-150	-125	-100	0	125	150	150	150	150	175	175
S15	-625	-250	-225	-200	-175	-150	-125	-100	0	100	120	140	140	150	150

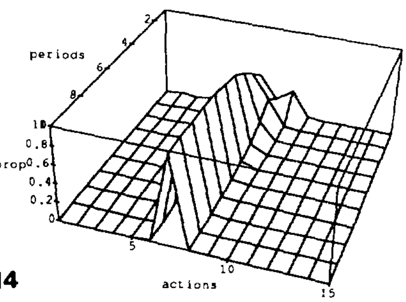
FIG. 12. Payoff matrix for Knott and Miller's experiment C.

- (S7,S7) is the pure strategy equilibrium.
- It is clear that both Cournot updating and fictitious play will eventually lead all the players to play the pure strategy Nash equilibrium. Indeed, when we ran fictitious play for 100 periods, all the mass converged to (S7,S7). Comparing the simulated frequencies from the two models with 10 periods of data, it is clear that the fictitious play hypothesis leads to much slower convergence to the pure strategy Nash equilibrium, and hence mimics the noisy data much better than the Cournot hypothesis for low values of ε.



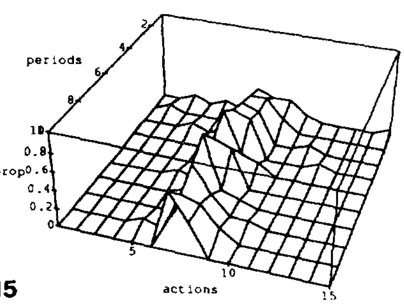
13

FIG. 13. Cournot simulated frequencies.



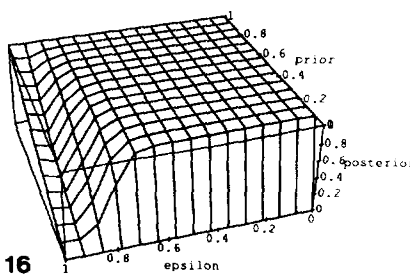
14

FIG. 14. Fict. play simulated frequencies.



15

FIG. 15. Observed frequencies.



16

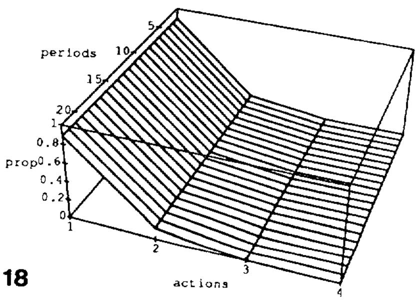
FIG. 16. Posterior on fictitious play.

Appendix 3. Experiment 3 of Cooper *et al.*[5]

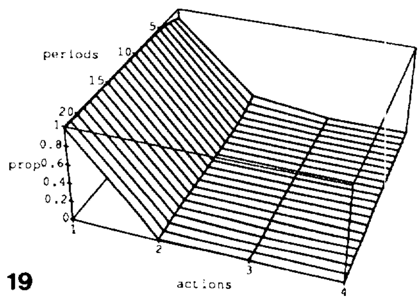
	S1	S2	S3
S1	350,350	350,250	1000,0
S2	250,350	550,550	0,0
S3	0,1000	0,0	600,600

FIG. 17. Payoff matrix for Cooper *et al.*'s experiment 3.

- (S1,S1) and (S2,S2) are pure strategy Nash equilibria.
- For almost all initial beliefs, S1 is a best response.
- If fictitious play is followed, everyone eventually plays S1.
- It is clear that the proportion of subjects playing S1 is not converging to 1 as fast as the fictitious play hypothesis predicts. Hence, for small values of ε, the Cournot hypothesis seems much better than the fictitious play hypothesis (since it always puts some mass on S2).

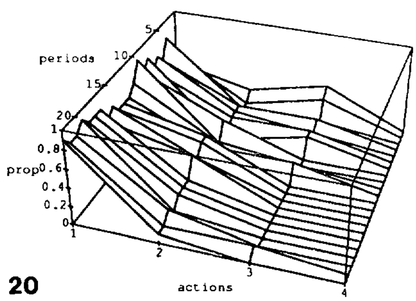


18

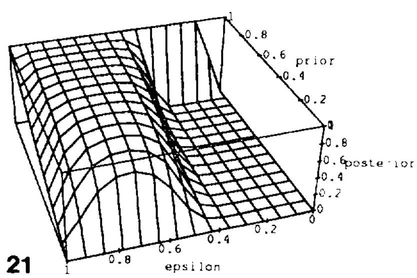


19

FIG. 18. Cournot simulated frequencies.
FIG. 19. Fictitious play simulated frequencies.



20



21

FIG. 20. Observed frequencies.
FIG. 21. Posterior on fictitious play.

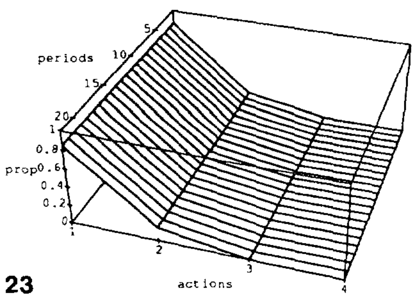
[5] Note that in Appendixes 3–8 there are only three strategies; the fourth strategy was only mandated by the limitations of our graphics package.

Appendix 4. Experiment 4 of Cooper *et al.*

	S1	S2	S3
S1	350,350	350,250	700,0
S2	250,350	550,550	0,0
S3	0,700	0,0	600,600

FIG. 22. Payoff matrix for Cooper *et al.*'s experiment 4.

- (S1,S1) and (S2,S2) are pure strategy Nash equilibria.
- For almost all initial beliefs, S1 is a best response.
- If fictitious play is followed, everyone eventually plays S1.
- The results of this experiment are not as sharp as for many of the others since the two hypotheses predict very similar behavior.



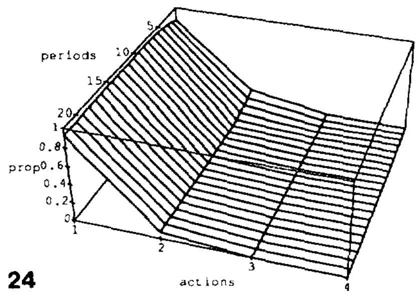
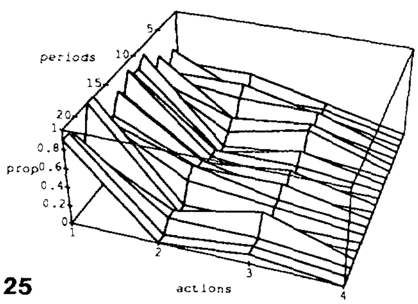


FIG. 23. Cournot simulated frequencies.
FIG. 24. Fictitious play simulated frequencies.



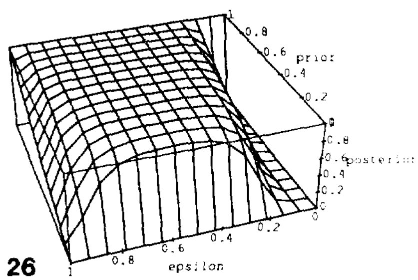


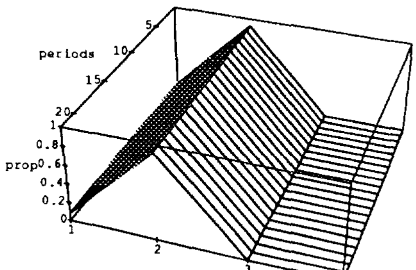
FIG. 25. Observed frequencies.
FIG. 26. Posterior on fictitious play.

APPENDIX 5. EXPERIMENT 5 OF COOPER *et al.*

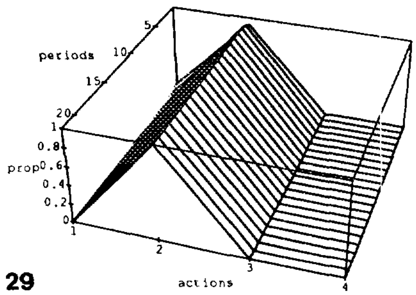
	S1	S2	S3
S1	350,350	350,250	700,0
S2	250,350	550,550	1000,0
S3	0,700	0,1000	600,600

FIG. 27. Payoff matrix for Cooper *et al.*'s experiment 5.

- (S1,S1) and (S2,S2) are pure strategy Nash equilibria.
- For almost all initial beliefs, S2 is a best response.
- If fictitious play is followed, everyone eventually plays S2.
- We can see that the Cournot hypothesis predicts that positive proportion of the players will continue to play S1, whereas the fictitious play hypothesis predicts that S1 will become extinct very quickly. Since the data have S1 die out rather rapidly, the fictitious play hypothesis fares very well at most values of ε.

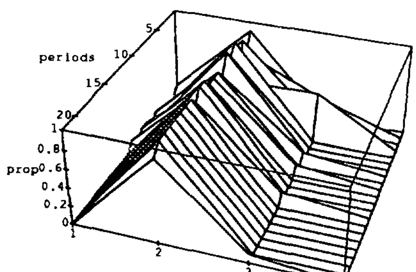


28

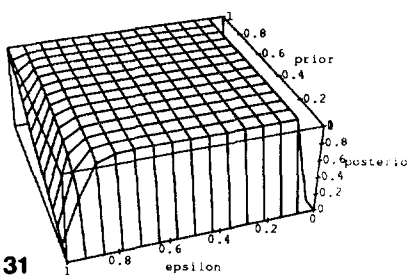


29

FIG. 28. Cournot simulated frequencies.
FIG. 29. Fictitious play simulated frequencies.



30



31

FIG. 30. Observed frequencies.
FIG. 31. Posterior on fictitious play.

	S1	S2	S3
S1	350,350	350,250	700,0
S2	250,350	550,550	650,0
S3	0,700	0,650	600,600

FIG. 32. Payoff matrix for Cooper *et al.*'s experiment 6.

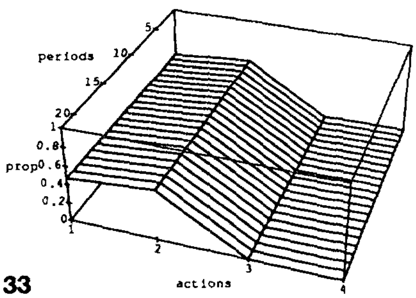
• (S1,S1) and (S2,S2) are pure strategy Nash equilibria.

• Approximately half the possible initial beliefs make S1 an optimal response and the other half make S2 a best response.

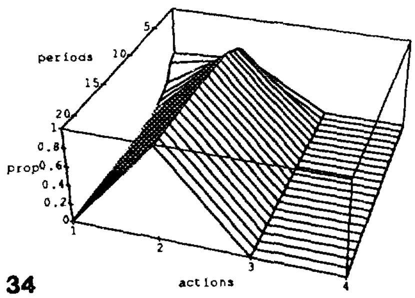
• However, if S3 is never played, then for almost all initial beliefs, S1 is a best response.

• If fictitious play is followed, and S3 is never played, everyone eventually plays S2.

• This is by far the most interesting experiment that we analyze since the dynamics predicted by the Cournot and fictitious play hypotheses are completely different. The behavior of the subjects is very similar to that predicted by the fictitious play hypothesis, and the result is a clear no-contest in favor of that hypothesis.

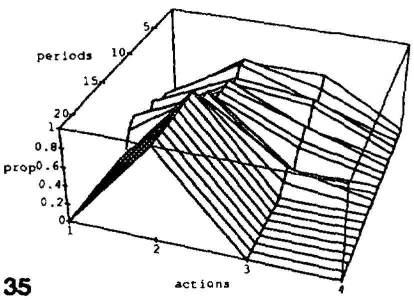


33

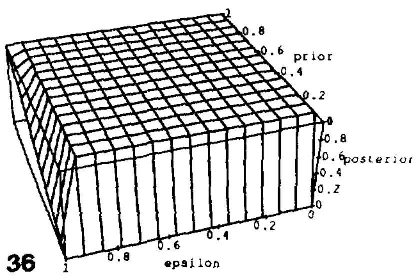


34

FIG. 33. Cournot simulated frequencies.
FIG. 34. Fictitious play simulated frequencies.



35



36

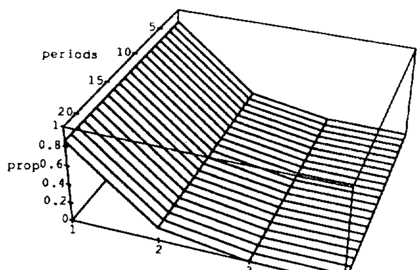
FIG. 35. Observed frequencies.
FIG. 36. Posterior on fictitious play.

Appendix 7. Experiment 7 of Cooper *et al.*

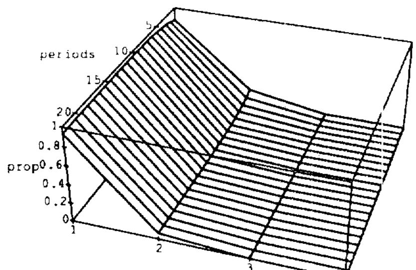
	S1	S2	S3
S1	350,350	350,250	700,0
S2	250,350	550,550	0,0
S3	0,700	0,0	500,500

Fig. 37. Payoff matrix for Cooper *et al.*'s experiment 7.

- (S1,S1) and (S2,S2) are pure strategy Nash equilibria.
- For almost all initial beliefs, S1 is a best response.
- If fictitious play is followed, everyone eventually plays S1.
- We see that both hypotheses predict for this experiment that most of the mass will converge to playing S1, whereas the data show that everybody eventually plays S2. Both hypotheses fare rather poorly for this experiment, but since the Cournot hypothesis predicts more mass at S2, it wins rather decisively at most values of ε.



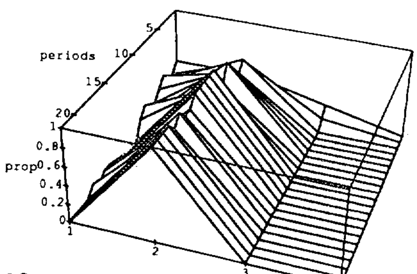
38



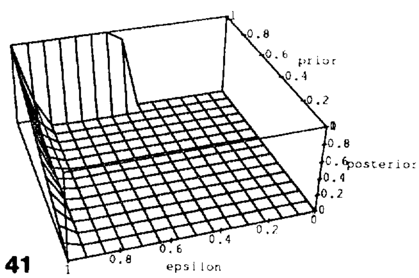
39

Fig. 38. Cournot simulated frequencies.

Fig. 39. Fictitious play simulated frequencies.



40



41

Fig. 40. Observed frequencies.

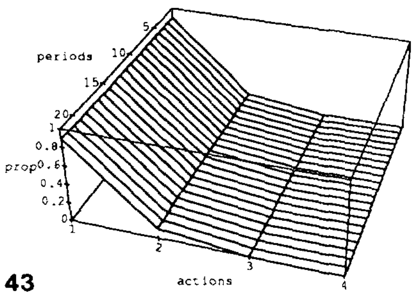
Fig. 41. Posterior on fictitious play.

Appendix 8. Experiment 8 of Cooper *et al.*

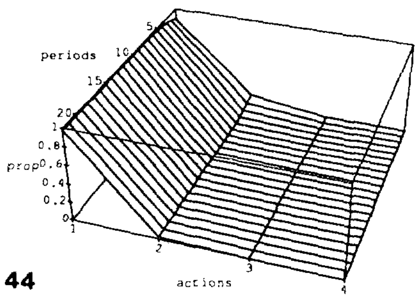
	S1	S2	S3
S1	350,350	350,250	1000,0
S2	250,350	550,550	0,0
S3	0,1000	0,0	500,500

FIG. 42. Payoff matrix for Cooper *et al.*'s experiment 8.

- (S1,S1) and (S2,S2) are pure strategy Nash equilibria.
- For almost all initial beliefs, S1 is a best response.
- If fictitious play is followed, everyone eventually plays S1.
- We see that both hypotheses predict for this experiment that most of the mass will converge to playing S1, whereas the data show that everybody eventually plays S2. Both hypotheses fare rather poorly for this experiment, but since the Cournot hypothesis predicts more mass at S2, it wins rather decisively at most values of ε.

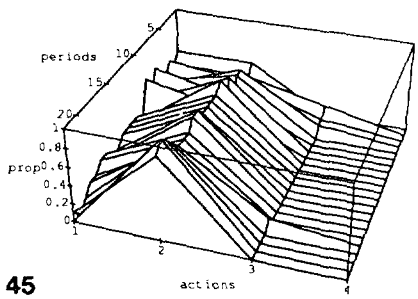


43

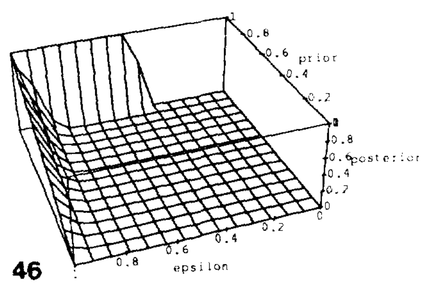


44

FIG. 43. Cournot simulated frequencies.
FIG. 44. Fictitious play simulated frequencies.



45



46

FIG. 45. Observed frequencies.
FIG. 46. Posterior on fictitious play.

REFERENCES

BROCK, W., MARIMON, R., RUST, J., AND SARGENT, T. (1988). "Informationally Decentralized Learning Algorithms for Finite-Player, Finite-Action Games of Incomplete Information," mimeograph, University of Minnesota.

BROWN, G. (1951). "Iterative Solutions of Games by Fictitious Play," in *Activity Analysis of Production and Allocation* (T-Koopmans, Ed.), pp. 377–380. New York: Wiley.

COOPER, R., DEJONG, D., FORSYTHE, R., AND ROSS, T. (1990). "Selection Criteria in Coordination Games: Some Experimental Results," *Amer. Econ. Rev.* **80,** 218–234.

EL-GAMAL, M. (1992). "The Extraction of Information from Multiple Point Estimates," mimeograph, California Institute of Technology.

EL-GAMAL, M., MCKELVEY, R., AND PALFREY, T. (1992). "A Sequential Bayesian Experimental Study of Learning in Games," *J. Am. Stat. Assoc.*, Forthcoming.

GENEST, C. AND ZIDEK, J. (1986). "Combining Probability Distributions: A Critique and an Annotated Bibliography," *Statist. Sci.* **1,** 114–148.

HOFFMAN, E., AND SPITZER, M. (1985). "Experimental Law and Economics," *Columbia Law Rev.* **85,** 991–1036.

KNOTT, J., AND MILLER, G. (1987). "The Dynamics of Convergence and Cooperation: Rational Choice and Adaptive Incrementalism." mimeograph Michigan State University.

MCKELVEY, R., AND PALFREY, T. (1992). "An Experimental Study of the Centipede Game," *Econometrica* **60,** 803–836.

MOULIN, H. (1985). *Game Theory for the Social Sciences*. New York: New York Univ. Press.

PLOTT, C. (1982). "Industrial Organization Theory and Experimental Economics," *J. Econ. Lit.* **20,** 1485–1527.

ROBINSON, J. (1951). "An Iterative Method for Solving a Game," *Ann. Math.* **54,** 296–301.

SHAPLEY, L. S. (1962). "On the Nonconvergence of Fictitious Play," memorandum RM-3026-PR, Rand Corporation.

SMITH, V. (1982). "Microeconomic Systems as an Experimental Science," *Amer. Econ. Rev.* **72,** 923–955.

WOLF, F. (Ed.) (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis*. Beverly Hills, CA: Sage.