

Rule Learning in Symmetric Normal-Form Games: Theory and Evidence¹

Dale O. Stahl

*Malcolm Forsman Centennial Professor, Department of Economics, University of Texas,
Austin, Texas 78712*

E-mail: stahl@eco.utexas.edu

Received November 13, 1997

An experiment, consisting of two 15-period runs with 5×5 games, was designed to test Stahl's [*International Journal of Game Theory* 28, 111–130 (1999)] model of boundedly rational behavioral rules and rule learning for symmetric normal-form games with unique symmetric Nash equilibria. A player begins with initial propensities on a class of evidence-based behavioral rules and, given experience over time, adjusts her propensities in proportion to the past performance of the rules. The experimental data provide significant support for rule learning and heterogeneity characterized by three modes. We also strongly reject "Nash learning" and "Cournot dynamics" in favor of rule learning. *Journal of Economic Literature* Classification Numbers: C72, C90, C51, C52. © 2000 Academic Press

Key Words: learning; rules; games; evidence

1. INTRODUCTION

Current research in game theory addresses the question of how players learn to play. At one end of the spectrum, we have the super-rational theories of Jordan (1991) and Kalai and Lehrer (1993), and at the other extreme, we have reinforcement learning models of Mookherjee and Sopher (1994, 1997), Roth and Erev (1995), and Erev and Roth (1998), in which players have only minimal intellect. The objectives of these inquiries are mixed, and include (i) a foundation for an improved equilibrium theory, and (ii) a realistic description of human behavior (especially in experimental games) whether or not that behavior is consistent with Nash

¹Partial funding of this research was provided by Grant SBR-9410501 from the National Science Foundation, but nothing herein reflects their views. Ray Battalio assisted in the design of the computer interface, Paul Wilson provided statistical guidance, Ernan Haruvy provided research assistance, and Benjamin Craig provided editing advice. I am also indebted to the editor in charge and the referees for constructive suggestions. However, all errors and omissions are the sole responsibility of the author.

equilibria. While the super-rational theories require unrealistically strong assumptions, the naive reinforcement models suffer from the opposite sin of assuming players are too limited in their reasoning powers.

The middle ground, in which players potentially have access to decision-making rules of varying sophistication and can learn which rules are appropriate in given situations, is probably much closer to the truth and, hence, would provide a far better descriptive and predictive model of human strategic behavior. If all that humans could learn about a game was which action to choose eventually, then nothing would be learned that could be reliably transferred to a different game. There is a clear advantage for humans to learn general rules which could be applied across a variety of games. The highest payoffs will go to those who choose actions that are optional with respect to the true probability distribution of play in the population for the next encounter, which entails being one step ahead of everyone else and, hence, requires flexibility in forecasting/decision rules. As teachers of economics and game theory, perhaps our greatest contribution is to introduce students to new ways to analyze problems and make decisions (such as Bayes rule). For our theories to continue to exclude such learning would be shortsighted.

Some steps in this direction have already been taken (Camerer and Ho, 1997, 1999; Cooper and Feltovich, 1996; Erev and Roth, 1998; Rapaport et al., 1997; Sonsino et al., 1998). The purpose of this paper is to take another step towards articulating and empirically testing such middle-ground theories.

Stahl (1999) put forth a theory of boundedly rational behavior characterized by “behavioral rules” and a theory of rule learning based on the performance of these rules. The general model is illustrated in Fig. 1. A *behavioral rule* is a function that maps from the available information (the game and any history of play) to the set of probability distributions on the actions available in the game. We also define a probability distribution (φ) over the space of behavioral rules. A random draw (from φ) selects a rule, which (given the available information) generates a probability distribution on the actions. A second random draw from this latter distribution consti-

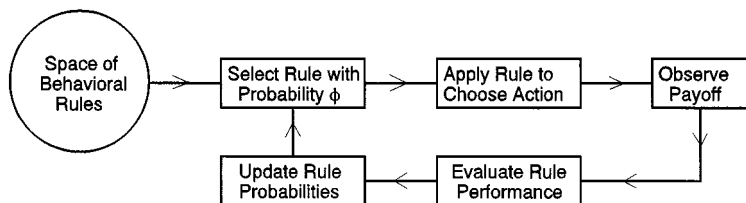


FIG. 1. Rule learning scheme.

tutes the chosen action. That action in combination with the choices of others in the population of players produces a payoff. In addition to knowledge of her payoff, the player receives information about the recent choices of the other players such as the recent empirical frequency of choices. This information is used to evaluate the performance of the behavioral rules. For example, the player could deduce the expected utility payoff of each rule against the recent past. Following the process of evaluating the rules, the probabilities (φ) to use the rules are updated. It is assumed that probabilities increase for rules that would have yielded higher payoffs in the recent past, and vice versa. This is the “Law of Effect” in psychology (Thorndike, 1898), and is also the logic behind evolutionary dynamics (Hofbauer and Sigmund, 1988).

For example, consider a simple space consisting of two rules: (1) “following the herd,” in which the current choice probabilities match the recent frequency distribution of choices, and (2) “Cournot best response” to the recent past. Let φ_0 denote the probability of using the herd rule, so $1 - \varphi_0$ is the probability of using the Cournot rule. A random draw picks one of the two rules—suppose the Cournot rule—which is then used to choose an action, which yields a payoff. Now suppose the player observes the population distribution of choices before having to make her next choice. The player can evaluate the hypothetical payoff she could have received had she used the other rule. If the following-the-herd rule would have yielded a greater payoff, then probability φ_0 increases, and vice versa. In the next period, a rule is selected according to the updated probabilities.

The first step in operationalizing this conceptual framework is to represent the huge space of potential behavioral rules. Our approach is to posit a small set of archetypal rules and a way of combining these rules to span a large space of plausible and empirically relevant rules. Since initial behavior is crucial to dynamic prediction, we need behavioral rules that are accurate for the initial period as well as later periods. Therefore, we build on the empirically successful level- n theory of bounded rationality for one-shot games developed by Stahl and Wilson (1994, 1995) [henceforth SW]. To extend this theory to multiple-period settings, we define history-dependent versions of the level- n rules. For example, the SW level-1 rule yields a noisy best response to the uniform distribution, and its multiperiod extension yields a noisy best response to the recent empirical frequency of play. One interpretation of this rule is that the expected utility of an action against the recent past is a kind of “evidence” for that action, with the player tending to choose the action with the most net favorable evidence. This interpretation suggests a large family of “evidence-based rules” which we develop in Section 2c. Evidence-based rules are generalizations of level- n rules that also admit Nash behavior, modified

versions of fictitious play and adaptive expectations, and mixtures of those rules. While not exhaustive, this family of rules is nonetheless very rich. Moreover, the interpretation of these rules as the result of “weighing the evidence” permits a natural interpretation of the learning dynamics as adjusting the weights on the evidence.

There are two kinds of learning present in the general model (recall Fig. 1). First, the players can learn in the sense of acquiring relevant data but sticking to the same behavioral rule. For instance, the two rules in the above example learn via data acquisition (the population distribution of choices). As another example, in a game against nature, Bayesian updating can be viewed as a fixed behavioral rule that learns via acquiring new data. We refer to this first sense of learning as “data learning.” Second, the players can learn by assessing the relative performance of the behavioral rules and switching to better performing rules. We refer to this second sense of learning as “rule learning.”

In Stahl (1996), a simplified version of this model was confronted with experimental “guessing game” data gathered by Nagel (1995), and the results were very encouraging. As an initial empirical test of the theory extended to normal-form games, Stahl (1999) used data gathered in an experiment that was primarily designed to test the robustness of the SW results, but since the experiment lasted two periods, the data could potentially provide evidence for learning. While two periods provide barely enough opportunity for learning by participants, the statistical test rejected the hypothesis of no rule learning. On the other hand, the no-rule-learning hypothesis could not be rejected for the majority of participants. This less than stellar individual learning result may reflect the difficulty of measuring learning over just two periods, rather than a falsification of the model. Clearly, more periods of data are needed for more conclusive testing of this model versus alternative models, which motivates the current experiment and paper.

In this paper, we make a small change in the Stahl (1999) specification of the evidence-based rules (Section 2b) which improves the fit, and then we estimate the parametric learning model for each participant of the experiment. While such a model with so many parameters has little practical value, it is quite useful for (i) addressing the general questions of learning unhampered by the false assumption of homogeneous participants, and (ii) characterizing the heterogeneity in the population. From our experimental data, we find that the hypothesis of no rule learning can be rejected in the aggregate and for a majority of the participants. We also find that there are three primary modes in the population distribution of parameters.

The paper is organized as follows. Section 2 presents the theory of rule learning and evidence-based boundedly rational rules. Section 3 presents

the experimental design and data. Section 4 presents the test results, and Section 5 summarizes and discusses the findings.

2. THEORY

We begin with a description of the game environment, then present the general theory of rule learning, and finally present the specific family of evidence-based rules and some operational specifications.

a. The Game Environment

Consider a finite, symmetric, two-player game $G \equiv (N, A, U)$ in normal form, where $N \equiv \{1, 2\}$ is the set of players, $A \equiv \{1, \dots, J\}$ is the set of actions available to each player, and U is the $J \times J$ matrix of expected utility payoffs for the row player, and U' , the transpose of U , is the payoff matrix for the column player. For notational convenience, let $p^0 \equiv (1/J, \dots, 1/J)'$ denote the uniform distribution over A .

We focus on single population settings in which each player observes the frequency distribution of the past action choices of all the other players in the population. In our opinion, this is a richer and more appropriate setting to study learning than settings in which each player is randomly matched with just one other player in each round and observes the choice of only that other player after that round. The latter setting is an exceedingly difficult setting in which to learn, because not only is the data process nonstationary, but it is also a poor sample of the population statistics until many periods of data are accumulated. Since the data learning task is much harder, sophisticated behavioral rules might not do better than simple rules until the players observe much more data, but experiments with many periods (say 100 or more) invite boredom effects and incentive problems. Further, in real-world learning, humans typically have ongoing access to a wealth of information about how other humans have behaved recently. While the pairwise random matching protocol is well suited to test one-shot equilibrium theory, it is not ideal for investigations of rule learning.

The empirical frequency of the *other* players' actions in period t will always be denoted by p^t , and this information is assumed to be available to each player.² The first period of play will be denoted by $t = 1$. It is also convenient to define $h^t \equiv \{p^0, \dots, p^{t-1}\}$, which is the *history* of other

²See Crawford (1994) and Cheung and Friedman (1997) on adaptive learning models for such situations.

players' choices up to period t with the novelty that p^0 is substituted for the null history. Thus, the information available to a player at the beginning of period t is $\Omega^t \equiv (G, h^t)$.

b. The General Theory of Rule Learning

A *behavioral rule* is a mapping from information Ω^t to $\Delta(A)$, the set of probability measures on the actions A . For the purposes of presenting the abstract model, let $\rho \in R$ denote a generic behavioral rule in a family of behavioral rules R ; $\rho(\Omega^t)$ is the mixed strategy generated by rule ρ given information Ω^t .

The second element in the general theory is a probability measure over the rules, $\varphi(\rho, t)$, which gives the probability of using rule ρ in period t . Given the family of behavioral rules R and probabilities φ , the induced probability distribution over actions for period t is the integral of the behavioral rules over the rule space: $\int_R \rho(\Omega^t) d\varphi(\rho, t)$.

It is convenient to specify the learning dynamics in terms of a state variable that is unrestricted in sign and magnitude. To this end, we define the state variable, $w(\rho, t)$, called the *log-propensity* for rule ρ in period t , such that the probability of using rule ρ in period t is

$$\varphi(\rho, t) \equiv \exp(w(\rho, t)) / \left[\int \exp(w(x, t)) dx \right]. \quad (1)$$

The last element of the general theory is the equation of motion for the log-propensities. The Law of Effect states that rules which perform well are more likely to be used in the future. This law is captured by the dynamic on log-propensities

$$w(\rho, t + 1) = \beta_0 w(\rho, t) + \beta_1 g(\rho, \Omega^{t+1}), \quad \text{for } t \geq 1, \quad (2)$$

where $g(\cdot)$ is the *reinforcement function* for rule ρ conditional on information $\Omega^{t+1} = (G, p^0, \dots, p^t)$. It is natural to assume that $g(\rho, \Omega^{t+1})$ is the expected utility that rule ρ would have generated in period t : $g(\rho, \Omega^{t+1}) = \rho(\Omega^t) U p^t$. Then, for small β_0 and large β_1 , past probabilities would be quickly swamped by new performance evidence, whereas for large β_0 and small β_1 , past probabilities would be strengthened despite new evidence.³

The reader may be more familiar with a one-parameter dynamic model in which $w(\rho, t + 1) = \beta_0 w(\rho, t) + (1 - \beta_0) g(\rho, \Omega^{t+1})$. While this one-parameter model has the desirable property of being asymptotically stable for $0 < \beta_0 < 1$, it has the shortcoming that we have no reason to believe

³Of course, "small" and "large" are relative terms that depend on the range of values taken by $w(\rho, \cdot) - w(\rho', \cdot)$ and $g(\rho, \cdot) - g(\rho', \cdot)$, which depend on the cardinal units used to measure expected utility.

that the scaling of expected utility implicit in $g(\cdot)$ is the correct scaling. To allow for this ignorance, the econometrician would multiply $g(\cdot)$ by a scalar ($\alpha > 0$) and estimate α . In our notation, $\beta_1 = (1 - \beta_0)\alpha$.

Given the rule space R and initial log-propensities $w(\cdot, 1)$, the law of motion, Eq. (2), completely determines the behavior of the system for all $t > 1$. The operational questions are (1) how to specify the rule space, and (2) how to specify the initial log-propensities.

An attractive feature of this general model (Fig. 1) is that it encompasses a wide variety of learning theories. For instance, to obtain replicator dynamics, we can simply let the rule space R be the set of J constant rules that always choose one unique action in A for all information states. Fictitious play and Cournot dynamics can be seen as very special cases in which R is a singleton rule which selects a best response to a belief that is a specific function of the history of play. Moreover, the general model can include these constant rules, best-response rules, and other rules.

In the next subsection, we introduce the family of evidence-based rules.

c. The Family of Evidence-Based Rules

We will first present the abstract concept of evidence-based rules, and then present the specific family of such rules that will be used in this paper.

An evidence-based behavioral rule arises from the notion that a player considers "evidence" for and against the available actions and tends to choose the action which has the most net favorable evidence given the available information. Suppose we hypothesize a finite number of kinds of evidence, indexed by $k \in \{0, 1, \dots, K\}$, and let $y_k \equiv \{y_{jk}\}_{j=1}^J$, a $J \times 1$ vector of real numbers, represent the k th kind of evidence: $y_{jk} > y_{j'k}$ means that the k th kind of evidence is more favorable for action j than action j' . For example, the vector of expected utilities against the recent past, Up^{t-1} , could be a kind of evidence. Let $Y \equiv \{y_0, \dots, y_k\}$ denote the $J \times (K + 1)$ matrix of evidence. It is important to understand that the evidence is based on the available information Ω^t , so we should write $Y(\Omega^t)$ to be perfectly clear.

Below, we will specify four kinds of evidence, but for now imagine a class of functions that operate on the available information Ω^t and produce a matrix of evidence. Further, suppose this class of functions is parameterized by θ , and let $Y(\Omega^t; \theta)$ denote the matrix of evidence given Ω^t and θ .

How will a player weigh such evidence? Let $\nu_k \geq 0$ denote a scalar weight associated with evidence y_k . We then define the weighted evidence vector

$$\bar{y}(\Omega^t; \nu, \theta) \equiv Y(\Omega^t; \theta)\nu, \tag{3}$$

where $\nu \equiv (\nu_0, \dots, \nu_K)'$. In other words, $\bar{y}(\Omega^t; \nu, \theta)$ is a $J \times 1$ vector of net evidence values, and we should expect that the strategy with the largest net evidence will be most likely to be chosen and vice versa.

There are many ways to go from such a weighted evidence measure to a probabilistic choice function. Like McKelvey and Palfrey (1995, 1998) and Anderson et al. (1997, 1999), we opt for the multinomial logit specification (McFadden, 1974) because of its computational advantages when it comes to empirical estimation. The interpretation is that the player assesses the weighted evidence with some error and/or adds idiosyncratic values to the expected utilities, and then chooses the action which from his/her perspective has the greatest net favorable evidence. Hence, the probability of choosing action j is

$$\hat{p}_j(\Omega^t; \nu, \theta) \equiv \exp[\bar{y}_j(\Omega^t; \nu, \theta)] / \sum_l \exp[\bar{y}_l(\Omega^t; \nu, \theta)]. \quad (4)$$

Note that Eq. (4) defines a mapping from Ω^t to $\Delta(A)$, and hence is a *behavior rule* as defined abstractly above. With a slight stretch of notation, we can associate the parameter vector (ν, θ) with a rule $\rho \in R$, in which case the family of behavioral rules R can be represented by (ν, θ) -space.

Our approach to representing the huge space of potential rules is to posit a small set of evidences which, when combined as in Eq. (3) and used in Eq. (4), spans a large space of plausible and empirically relevant rules. Since the concept of evidence-based rules can be applied to one-shot behavior as well as dynamic behavior, the predicted behavior in the first period of play should be consistent with the predicted behavior of boundedly rational theories for one-shot games.

In one-shot games, SW found substantial support for level- n thinking. A level- n player believes that all others are level- $(n - 1)$ players who believe... are level-1 players who believe that all others are level-0 types who choose randomly. In a repeated guessing game, Nagel (1995) and Stahl (1996a) found substantial support for similar thinking and rule learning. These types were investigated because they correspond loosely to the iterated levels of rationalizability (Bernheim, 1984 and Pearce, 1984; see also Stahl, 1993), and Binmore's (1987) idea of truncated eductive reasoning. It is also reasonable to ask whether a player who seems to use a level-1 rule in a one-shot game would continue to do so in subsequent plays of that game, especially if, say, the level-2 rule would have produced higher payoffs. The evidence-based rules described below are an extension of the level- n theory to a dynamic situation with learning opportunities.

We now posit four kinds of evidence such that each kind corresponds to one of the SW level- n types. The first kind of evidence comes from a "null" model of the other players. The null model provides no reason for

the other players to choose any particular strategy, so for the first period of play, by virtue of insufficient reason, the belief is that all actions are equally likely. The expected utility payoff to each available action conditional on the null model is $y_1(\Omega^1) \equiv Up^0$. We can interpret y_{j1} as evidence in favor of action j stemming from the null model in the initial period of play (with no prior history). Note that if this is the only kind of evidence given any weight, then the probabilistic choice function, Eq. (4), would result in a (possibly noisy) best response to p^0 , which is precisely the level-1 archetype of SW.

For later periods ($t > 1$), the players have empirical data about the past choices of the other players. It is reasonable for a player to use simple distributed-lag forecasting: $(1 - \theta)p^0 + \theta p^1$ for period 2 with $\theta \in [0, 1]$. Letting $q^t(\theta)$ denote the forecast for period t and defining $q^0(\theta) \equiv p^0$, the following forecasting equation would then apply for all $t \geq 1$:

$$q^t(\theta) \equiv (1 - \theta)q^{t-1}(\theta) + \theta p^{t-1}.^4 \tag{5}$$

The expected utility payoff given this forecast is $y_1(\Omega^t; \theta) \equiv Uq^t(\theta)$. We call $y_{j1}(\Omega^t; \theta)$ the “level-1” evidence in favor of action j stemming from the null model and prior history.

The second kind of evidence is based on the SW “level-2” player who believes all other players are level-1 players. We define the archetypal level-2 player as one whose belief about the other players is $b(q^t(\theta))$, which puts equal probability on best responses to $q^t(\theta)$ and zero probability on all other actions.⁵ The expected utility conditional on this belief is $y_2(\Omega^t; \theta) \equiv Ub(q^t(\theta))$, and we call $y_{j2}(\Omega^t; \theta)$ the “level-2” evidence in favor of action j . Note that if this is the only kind of evidence given any weight, then the probabilistic choice function, Eq. (4), would result in a noisy best response to the best response to the uniform prior (in period 1), which is precisely the level-2 archetype of SW.

In testing alternative theories, it is very useful to have an encompassing model; thus, we want to incorporate Nash equilibrium theory within the model. Letting p^{NE} denote a Nash equilibrium of G , $y_3 \equiv Up^{\text{NE}}$ provides yet another kind of evidence on the available actions. Note that this kind of evidence is not well defined for games with multiple Nash equilibria. Since the resolution of this multiplicity problem is beyond the scope of this paper, our experimental design uses only games with a unique Nash

⁴An alternative specification could be to set $\theta = 1/t$, which would generate the empirical frequency as the forecast as in fictitious play. However, a declining weight on the most recent observation may not be empirically true. Our specification allows θ to be learned, and it is mathematically possible that the learned value of θ could decline as $1/t$. Thus, fictitious play is a possible dynamic path.

⁵This is the limit of a logistic best response as the precision goes to infinity.

equilibrium. Again, if this is the only kind of evidence given any weight, then the probabilistic choice function, Eq. (4), would result in a noisy best response to the Nash prior, which is precisely the naive Nash archetype of SW.

Finally, we would like to represent behavior that is uniformly random in the first period (i.e., SW level-0) and “follows the herd” in subsequent periods. Following the herd does not mean exactly replicating the most recent past, but rather following the past with perhaps some inertia as represented by $q^t(\theta)$. Accordingly, we define $y_0(\Omega^t; \theta) = \ln[q^t(\theta)]$, so if only this evidence were given positive weight, then the logit formula, Eq. (4), would reproduce $q^t(\theta)$ as the probabilistic choice function.

Note that a single scalar value of θ is hypothesized for all its uses in $y_k(\Omega^t; \theta)$, $k = 0, 1$, and 2. This simplifying assumption can be defended by arguing that a player begins with a tendency to follow the herd (i.e., evidence $y_0(\Omega^t; \theta)$), and then perhaps reasons further. If the player reasons one step further, then the player projects his/her own $y_0(\Omega^t; \theta)$ onto the other player, thereby generating evidence $y_1(\Omega^t; \theta)$ with the same θ . If the player reasons one step further, then the player projects his/her own $y_1(\Omega^t; \theta)$ onto the other player, thereby generating evidence $y_2(\Omega^t; \theta)$ with the same θ .⁶

In summary, we have defined four kinds of evidence: herd $y_0(\Omega^t; \theta)$, level-1 $y_1(\Omega^t; \theta)$, level-2 $y_2(\Omega^t; \theta)$, and Nash y_3 , which we denote compactly by a $J \times 4$ matrix $Y(\Omega^t; \theta)$. Given the weight vector $\nu = (\nu_0, \dots, \nu_3)'$, the weighted evidence vector is $\bar{y}(\Omega^t; \theta, \nu) = Y(\Omega^t; \theta)\nu$, and the probabilistic choice function is given by Eq. (4). The archetypal level-1 rule corresponds to the weight vector $(0, \nu_1, 0, 0)$, etc. But there is no reason to restrict the rule space R to only these four archetypal rules. By letting ν be a continuous variable, we generate a five-dimensional family of behavior rules characterized by $(\nu_0, \dots, \nu_3, \theta)$. Each point in this rule space defines a unique behavioral rule resulting from a combination of the four kinds of evidences. Rule learning within this family of evidence-based rules amounts to adjusting the weights on these four kinds of evidence.

For another interpretation, we can rewrite the combination of level-1, level-2, and Nash evidence as $vU[\tilde{\nu}_1 q(\theta) + \tilde{\nu}_2 b(q(\theta)) + \tilde{\nu}_3 p^{\text{NE}}]$, where $\tilde{\nu}_k \equiv \nu_k/v$ and $v \equiv \sum_{k=1}^3 \nu_k$. Thus, the weighted evidence entails an implicit mixture of the primitive priors $\{q(\theta), b(q(\theta)), \text{and } p^{\text{NE}}\}$ corresponding to the above three models of other players. Each $\tilde{\nu}_k$ can be interpreted as the player's point estimate of the proportion of other players who can

⁶In Stahl (1999), level-1 evidence was defined isomorphically, but level-2 evidence and herd behavior implicitly assumed $\theta = 1$. In addition to the foregoing defense of having θ affect all level- n rules ($n = 0, 1, 2$), preliminary investigations found a substantial improvement in the fit to the experimental data.

be characterized by each of these archetypal models. Under this interpretation, adjusting the ν weights can be viewed as learning about those population proportions.

d. Initial Log-Propensities and Transference

An individual player starts with initial log-propensities to use these rules in period 1, which we denote $w(\nu, \theta, 1)$, and the log-propensities evolve according to the law of motion, Eq. (2). It remains to specify this initial log-propensity function. A natural and parsimonious approach is to specify $w(\cdot, 1)$ as a normal distribution in rule space with a mean $(\bar{\nu}, \bar{\theta})$ and standard deviation σ to be estimated from the data. While simplistic, it is a proper prior on the rule space and is compatible (when σ is relatively small) with SW who found that the vast majority of participants appeared to use one rule for a variety of one-shot 3×3 symmetric games.

Anticipating the experiment which will consist of two “runs” of 15 periods each with one game in the first run and a different game in the second run, we must also specify the initial log-propensities for the first period of the second run (i.e., period 16). One approach would be to specify $w(\cdot, 16)$ as another normal distribution in rule space with a potentially independent mean and standard deviation. However, this approach would involve the addition of five more parameters. Another approach would be to assume that one initial log-propensity applies to both periods, as if nothing that was learned during the first run is transferred to the second run. A third alternative would be to assume complete transference; i.e., the log-propensity for the first period of the second run is the same as it would be if there had been a sixteenth period of the first run. We opt for a convex combination of the second and third alternatives which requires only one additional parameter, i.e.,

$$w(\nu, \theta, 16) = (1 - \tau)w(\nu, \theta, 1) + \tau w(\nu, \theta, 15+), \quad (6)$$

where “15+” indicates the update after period 15 of the first run, and τ is the *transference* parameter. If $\tau = 0$, there is no transference, so period 16 has the same initial log-propensity as period 1; and if $\tau = 1$, there is complete transference, so the first period of the second run has the log-propensity that would prevail if it were period 16 of the first run (with no change of game). Note that this approach can be extended to multiple runs with different games with no need for additional parameters (assuming a constant transference τ).

e. The Likelihood Function

Let $s^h \equiv (s^{h1}, \dots, s^{hT}) \in \{1, 2, 3, 4, 5\}^T$ denote the choices of participant h for an experiment with T periods. The theoretical model put forth to

explain these choices involves nine parameters: $\beta \equiv (\bar{\nu}_0, \bar{\nu}_1, \bar{\nu}_2, \bar{\nu}_3, \bar{\theta}, \sigma, \beta_0, \beta_1, \tau)$. The first five parameters $(\bar{\nu}_0, \dots, \bar{\nu}_3, \bar{\theta})$ represent the mean of the participant's initial log-propensity $w(\cdot, 1)$, and σ is the standard deviation of that log-propensity; β_0 and β_1 are the learning parameters of Eq. (2); and τ is the transference parameter in Eq. (6) for the initial log-propensity of the second run.

Letting $\varphi^h(\nu, \theta, t | \beta)$ denote individual h 's probability to use rule (ν, θ) in period t (with information Ω^t) given the nine-parameter vector β , the resulting probability of choosing action j is

$$p_j^{ht}(\beta) = \int \hat{p}_j(\Omega^t; \nu, \theta) \varphi^h(\nu, \theta, t | \beta) d(\nu, \theta). \quad (7)$$

Then the log-likelihood of s^h conditional on β is

$$\text{LL}^h(\beta) \equiv \sum_{t=1}^T \ln[p_{s^h}^{ht}(\beta)]. \quad (8)$$

The computation of Eq. (7) entails numerical integration on a grid; details are available from the author upon request.

f. Limit Behavior

Observe that if behavior converges, the limit behavior will be a Nash equilibrium. To see this, suppose to the contrary that empirical frequency converges to a non-Nash limit point. But then the rule which puts infinite weight on level-1 evidence and no weight on any other evidence, since it will generate the best response to the empirical distribution, will perform at least as well as any other rule, and hence will increase in likelihood, thereby moving the empirical frequency away from the non-Nash limit point—a contradiction. However, if there is an upper bound on the evidence weights ν , then the limit behavior will be a quantal response equilibrium.

Further, observe that (since in symmetric games when the distribution is Nash, then level-1, level-2, and herd behavior are all the same) all the rules will perform equally well. Hence, if behavior converges, “rule learning” will become negligible as time advances.⁷ Thus, rule learning will be important only in cases for which behavior does not converge or in the early periods of convergent cases. This fact justifies our focus on rule learning in the “short run” and our experiment design.

⁷This is reminiscent of the principle, from Stahl (1993), that “being right is just as good as being smart.”

3. EXPERIMENT DESIGN AND DATA

In contrast to the experimental design of Stahl (1999), which presented participants with 15 games in each of two periods, the current design presents participants with one game for 15 periods. The former experiment was a follow-up to SW to provide more observations per participant facilitating participant-by-participant estimation of the level- n model, and only secondarily to obtain data to test the learning model. To provide a fuller test of the learning model, we need many more periods of data, but to identify different rules we need strategic variety. That is, we need either a variety of simple games (such as the 3×3 games of the former experiments) or more complicated games. The second route is better suited for gathering multiperiod data within the time constraints of an experiment.

With four archetypal rules ($\nu_k > 0$ for just one k), we reasoned that an appropriate 5×5 game should be able to elicit distinct behavior for these four archetypal rules as well as “interior” rules. We used a computer to search for symmetric games which satisfy the following criteria:

- (1) the first strategy is a strict Nash equilibrium and there are no other symmetric Nash equilibria;
- (2) the second strategy is the best response to uniform;
- (3) the third strategy is the twice-iterated best response to uniform (i.e., best response to the second strategy);
- (4) the fourth strategy is the response of the “worldly” type (Stahl and Wilson, 1995)⁸; and
- (5) the fifth strategy is either strictly dominated or the maximax choice.

Far from giving the evidence-based rule learning theory its “best shot,” these criteria create the greatest possibility of falsification: if different rules would produce different observable behavior and the no-rule-learning hypothesis cannot be rejected, then that would be much more damaging than a nonrejection of the null with data in which the different rules do not produce different observable behavior.

In addition, payoffs were confined to be integers from 0 to 99. Searching over billions of randomly generated 5×5 matrices yielded about a dozen candidate games.⁹ From this group we eliminated games for which the

⁸The “worldly” type is roughly equivalent to an evidence-based rule with weights $\nu_1 = 0.015$, $\nu_2 = 0.035$, $\nu_3 = 0.05$. Note that ν^0 is irrelevant for the first-period prediction.

⁹We are not counting the equivalent matrices obtained by permuting the rows and columns. Confining payoffs to be integers leaves $100^{25} = 10^{50}$ possible matrices, so random generation ensures statistically adequate sampling of all possible matrices.

Nash equilibrium was very focal because it had a large payoff (80 or more), or very antifocal because it had a very small payoff (10 or less). Following this selection, the rows and columns were permuted so a particular behavior rule would not be associated with a particular strategy in all games.

The payoffs for the “row player” in the four games we selected are shown in Fig. 2. Payoffs are in probability units for a fixed prize of \$2.00 per game.¹⁰ The labels ne, b1, b2, wd, dm, and mx in Fig. 2 denote respectively the strategies that satisfy criteria (1)–(5) above.

Each game (“Decision Matrix”) was presented on the computer screen (see Fig. 3). The participant made a choice of a pure strategy by clicking on a row of the matrix, which then became highlighted. In addition, the participant could enter a hypothesis about the choices of the other players, and cause the computer to calculate hypothetical earnings, which were then displayed on the screen. This feature was provided for two reasons. First, three of the kinds of evidence, level-1, level-2, and Nash, require calculation of the expected utility payoff against a prior, and this is difficult in a 5×5 game. Second, reducing calculation noise will sharpen the theory predictions, thereby increasing its falsifiability.

After each period, each participant was matched with every other participant and received the average payoff. Following this, each participant was shown his/her payoff and the aggregate choices of the *other* participants. The most recent period’s data was displayed on the main screen, but participants could access the entire past record with the click of a mouse button. The lotteries that determined final monetary payoffs were conducted following the completion of both runs. Delaying the lotteries avoids the possibility that lottery outcomes bias the reinforcement of rules. The average payment per participant was \$28.00 for a $2\frac{1}{2}$ -hour session.

At the beginning of the experiment, participants were given extensive instructions (both on-screen and read aloud) regarding the computer interface, the on-screen calculator, and the manner in which their choices would determine their payoffs. The experiment script is available from the author upon request. All participants passed a quiz on these matters.

¹⁰ Binary lotteries might not succeed in inducing risk neutrality. On the other hand, binary lottery payoffs also admit nonexpected utility formulations of preferences as long as people are probabilistically sophisticated (i.e., satisfy the compound lottery reduction axiom) and utility is monotonic in the probability of winning (Machina and Schmeidler, 1992). Under these strong conditions, “expected utility” can be replaced everywhere in this paper by “expected probability of winning.”

Game I:

19	43	96	85	85	b1
28	62	88	74	24	ne
67	21	38	48	38	b2
40	58	0	15	92	wd
16	15	86	99	79	mx

Game II:

68	10	76	33	75	wd
73	4	59	0	8	dm
3	92	16	15	99	b2
86	54	25	41	6	ne
72	98	92	8	52	b1

Game III:

2	31	0	99	6	mx
6	10	97	40	24	b2
98	96	38	48	19	b1
42	40	80	51	48	wd
97	46	5	68	49	ne

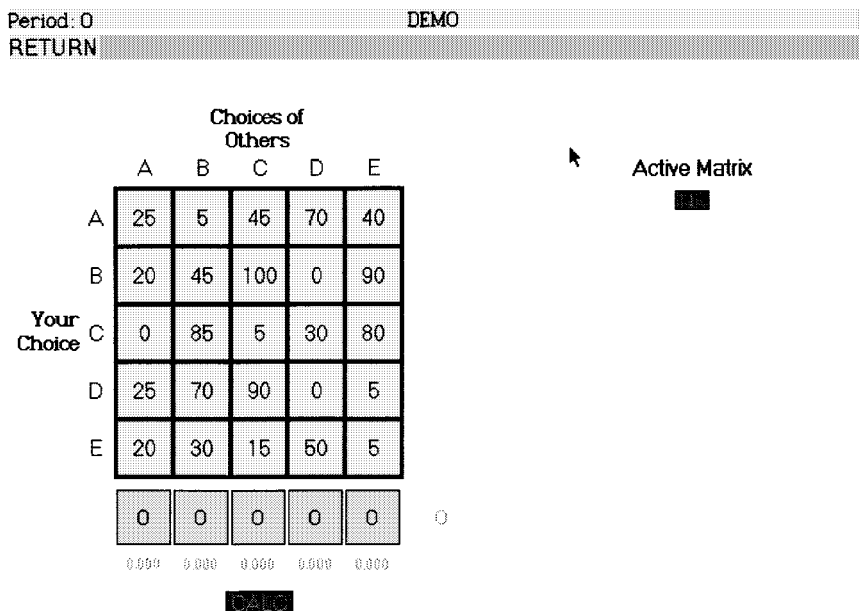
Game IV:

22	79	35	56	75	wd
22	38	78	55	99	b1
27	58	1	11	0	dm
70	1	34	59	37	ne
56	84	60	23	2	b2

KEY:

ne = Nash equilibrium
 b1 = best response to uniform
 b2 = best response to b1
 wd = SW "worldly" choice
 dm = dominated
 mx = maximax

FIG. 2. The four games.



You have 1 matrix left.
FIG. 3. Computer screen.

The experiment consisted of four sessions of 22, 23, 24, and 22 participants.¹¹ The participants were predominantly upper-division undergraduate students and some graduate students attending the first and second 1995 summer sessions at the University of Texas. Each session consisted of two runs of 15 periods each. In the first run, one of the four games was played for 15 periods, and in the second run, another game was played for 15 periods. The specific order of presentation in the four sessions was (II, I), (III, IV), (IV, III), and (I, II) respectively. In each run, the first five

¹¹In two sessions, there was one participant who failed to make a choice in one period of one of the runs. In subsequent periods of that run, the computer excluded that participant's choices when calculating payoffs and the history for the other participants. Consequently, we did not estimate a likelihood function for these participants and did not include them in the above totals. However, when calculating the likelihood function for all other participants, the history of play used in this calculation was the exact history seen on the computer screen, which included the errant participant's choices before his or her mistake but not afterwards.

periods were 3 minutes each, and the last ten periods were 2 minutes each.¹²

Figures 4a–h display the empirical distribution of choices for each of the runs. In two runs, behavior converged to the unique Nash equilibrium (S627r2 and S810r2)¹³; in addition, in S629r2, the Nash equilibrium became the modal response. In all runs except S629r2, the modal response in the first period was the best response to the uniform prior (level-1 behavior). For the two runs of game I (S627r1 and S815r2), there is a curious persistence of “cooperative” behavior in that a large portion of participants chose the top row (A), which (if everyone chose it) would yield the symmetric Pareto dominant outcome. While the game is technically a repeated game, it is unusual to observe cooperative behavior in a group this large. Further data analysis revealed that row A was often a best (or nearly best) response to the past frequency distribution; indeed, there are quantal response equilibria in the neighborhood of this frequency distribution. Subsequent pilot experiments with different matching protocols designed to frustrate intertemporal effects produced similar data, thereby convincing us that the data used in this paper are not the result of supergame strategies.¹⁴

4. RESULTS

We divide the results into two sections. The first section deals with hypothesis tests, and the second section analyzes the parameter estimates for the individual participants. Computational details are available from the author.

a. Hypothesis Tests

Maximizing the log-likelihood (LL) function for each participant, and then summing over the 91 participants, the aggregate maximized LL is -1896.44 . We report in this subsection selected results on hypotheses that

¹²These durations were chosen after reviewing the responses to a debriefing questionnaire following a pilot experiment. Fewer than 10% of the pilot participants indicated that they wanted more time. Nonetheless, it is possible that the time constraint could have been binding for some participants and thereby inhibit their ability to use or learn more sophisticated rules.

¹³“S_{xxxr}y” refers to Session *xxx* (date) and run *y* (1 or 2).

¹⁴The alternative matching protocols included (i) matching each participant with a different random sample of 25% of the other participants each period, and (ii) ordering the participants around a virtual circle and having the effect of one’s choice move counterclockwise and only part way around the circle. Further, several intertemporal punishment strategies were hypothesized, tested, and rejected.

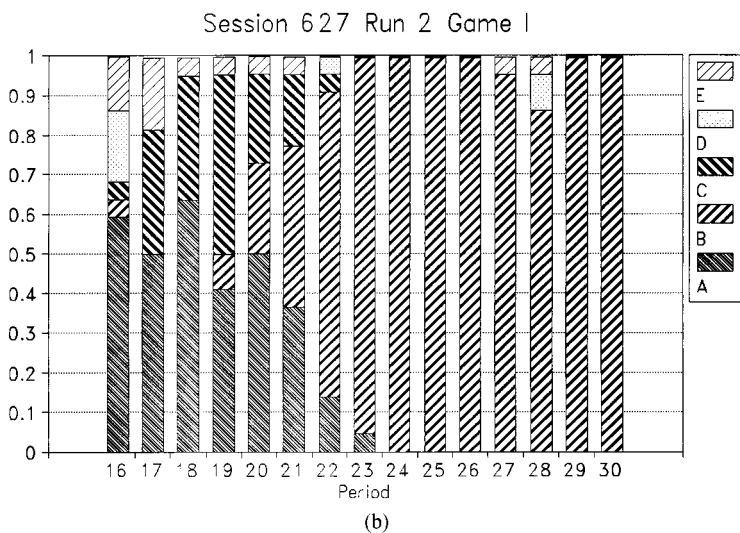
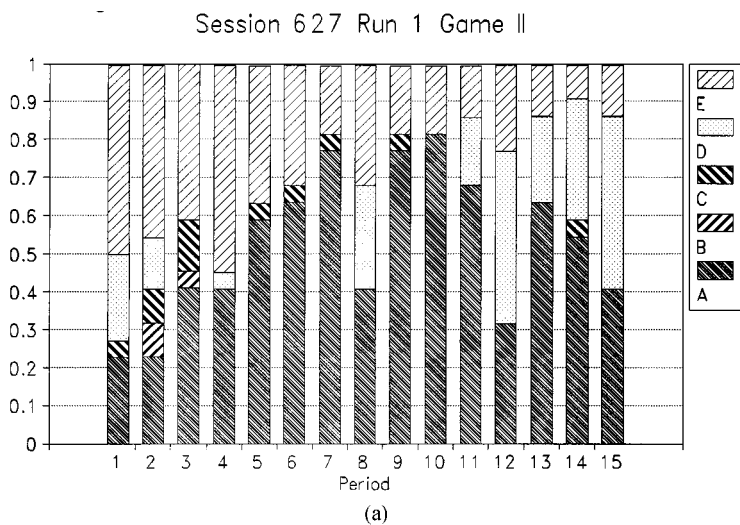
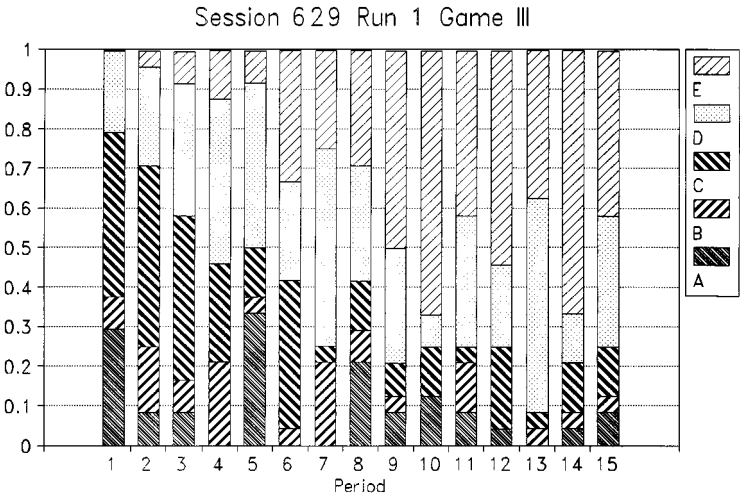
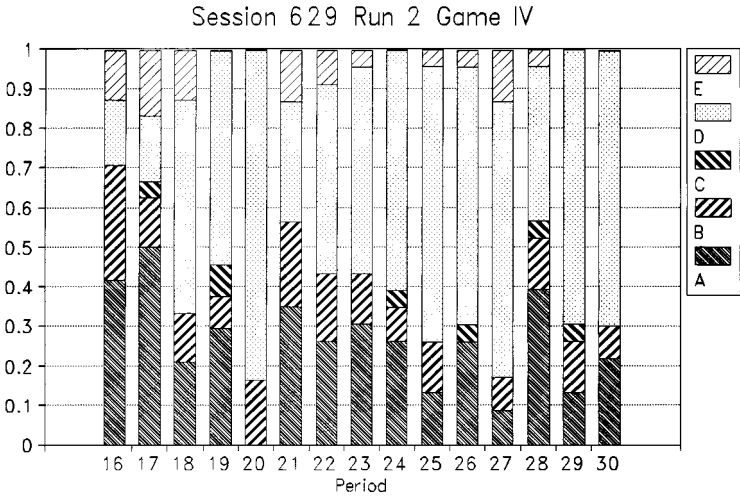


FIG. 4. Distribution of choices.



(c)



(d)

FIG. 4. —Continued.

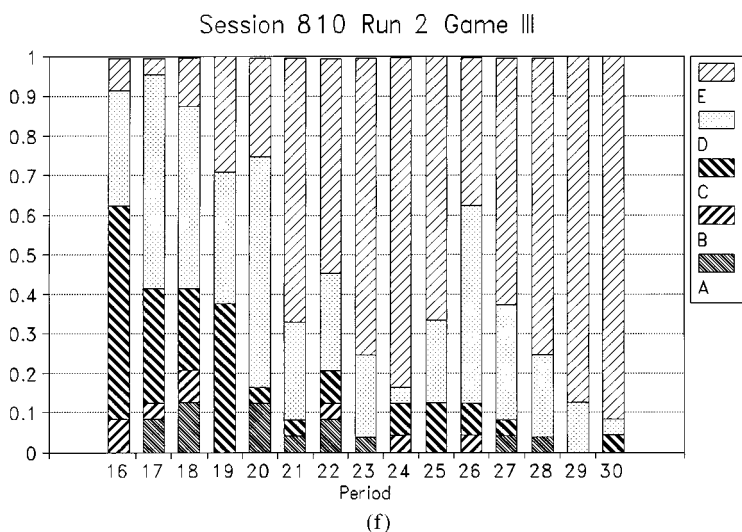
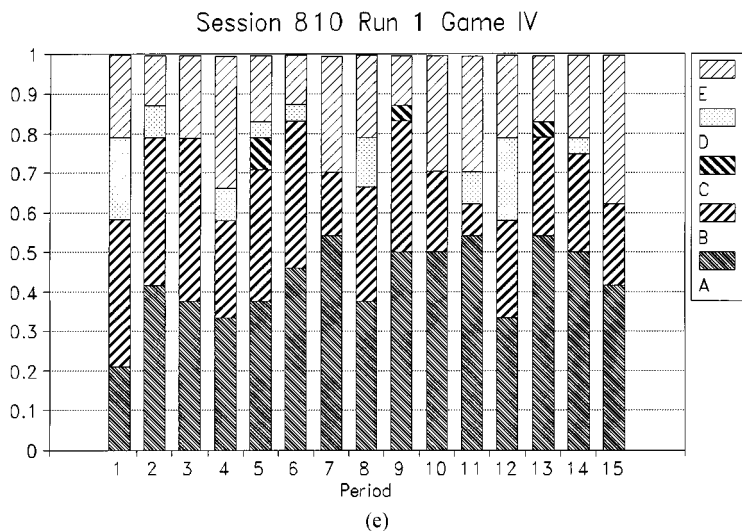


FIG. 4. —Continued.

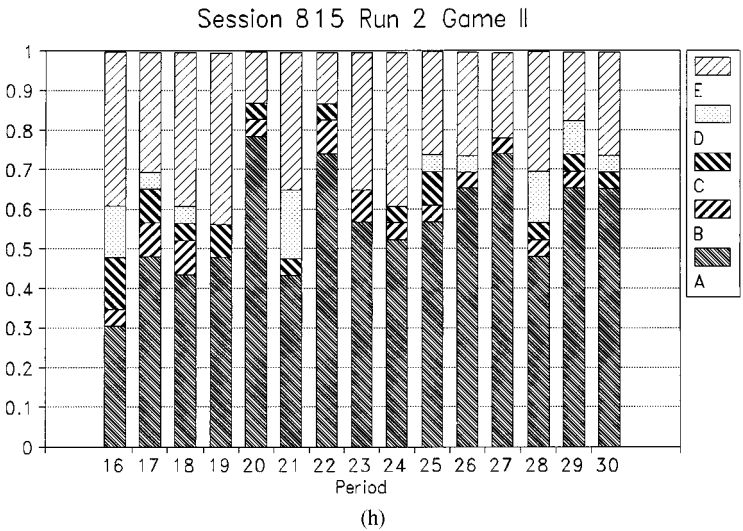
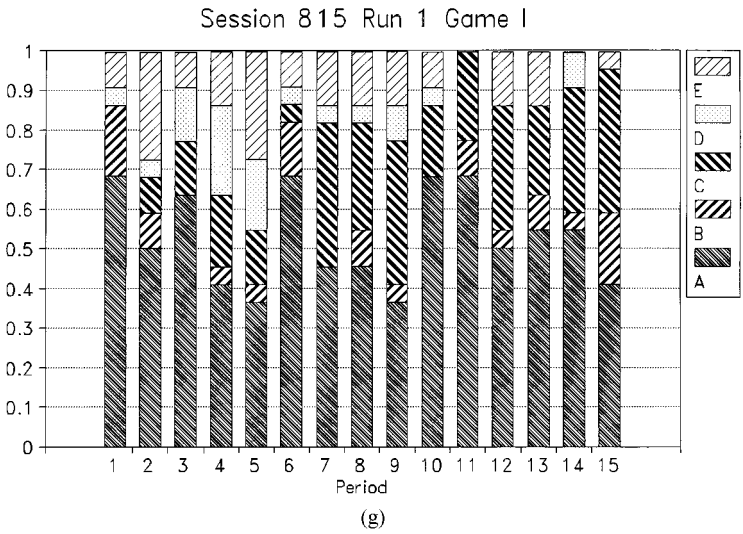


FIG. 4. —Continued.

TABLE I
Aggregate Hypothesis Tests

Hypothesis	LL	d.f.	p <
Noise	-4393.75	819	10^{-587}
Nash	-3917.14	637	10^{-486}
No Rule Lrn	-2303.01	273	10^{-54}
Cournot	-2829.30	637	10^{-120}
Homogeneous	-3060.88	810	10^{-146}
$\bar{\nu}_0 = 0$	-2014.41	91	10^{-14}
$\bar{\nu}_1 = 0$	-2315.34	91	10^{-120}
$\bar{\nu}_2 = 0$	-2022.61	91	10^{-16}
$\bar{\nu}_3 = 0$	-1987.33	91	10^{-7}
$\bar{\theta} = 0$	-2167.25	91	10^{-64}
$\beta_1 = 0$	-2151.30	91	10^{-58}
$\tau = 0$	-2125.14	91	10^{-49}
$\beta_0 \leq 1$	-2012.66	91	10^{-13}
Unrestricted	-1896.44	0	

impose restrictions on the parameters which lower the maximized LL function. These results are summarized in Table I.

RESULT 1. The model is informative and fits the data much better than Nash equilibrium theory.

In multiple regression analysis, the F-test for the whole model serves as a test of whether the model is informative (i.e., fits the data better than pure noise). In the context of maximum likelihood estimation, the equivalent test uses the null hypothesis of uniformly random play in all games and periods. This null hypothesis yields an aggregate LL value of -4393.75 ($= 30 \times 91 \times \ln(5)$). Since the model is capable of generating uniformly random play (by restricting $\sigma = 0$, $\bar{\nu} = 0$, $\beta_0 = 1$, and $\beta_1 = 0$), the null is a nested hypothesis. Therefore, twice the log-likelihood difference is distributed χ^2 with 819 ($= 9 \times 91$) degrees of freedom, which has a p-value of less than 10^{-587} ; thus, the model is clearly informative.

As another benchmark, we also consider the Nash equilibrium model. Of course, in its pure form, it is incompatible with the data because participants often make non-Nash choices. It is more interesting to consider the Nash model extended to include errors. Observe that by setting $\nu_k = 0$, $\forall k \neq 3$, Eqs. (3)–(4) define a Nash-based probabilistic choice function, with the interpretation of ν_3 as the precision of the participant's expected utility calculation given the Nash belief. The hypothesis that participants make their choices according to this error-prone Nash model can be represented by a parameter restriction on our full model: namely, $\sigma = 0.25$, $\bar{\nu}_k = 0$ for $k \neq 3$, $\beta_0 = 1$, and $\beta_1 = 0$. In other words, this

Nash-based model is nested within our full model as a restriction on seven parameters.¹⁵ For each participant, we found the $(\bar{v}_3, \bar{\theta})$ values that maximized the log-likelihood of his/her choices. The sum over all 91 participants of these maximized LL values was -3917.14 . Compared with the totally random prediction (-4393.75), this is a significant improvement ($p < 10^{-104}$ with 2×91 degrees of freedom). However, the full model (-1896.44) is a very significant improvement over this Nash model ($p < 10^{-486}$ with 7×91 degrees of freedom). In other words, even after adjusting for the large number of parameters in the full model, the full model is astronomically more likely to have generated the data than the Nash-based model. An enhanced Nash model with learning is also rejected.

RESULT 2. We can reject the hypothesis of no rule learning.

By rule learning we mean that the probabilities (φ) to use rules change from period to period in accordance with the law of motion, Eq. (2). The null hypothesis of no rule learning implies the restriction that $\beta_0 = 1$ and $\beta_1 = 0$, and hence is a nested hypothesis. (Note that since τ becomes irrelevant, these restrictions entail three degrees of freedom.) When we impose these restrictions, the aggregate maximized LL decreases to -2303.01 . Twice the difference is distributed χ^2 with 273 ($= 3 \times 91$) degrees of freedom and has a p-value less than 10^{-54} . Thus, we can strongly reject the hypothesis of no rule learning.

On an individual basis, for 50 of 91 (54.9%) of the participants, we can reject the no-rule-learning hypothesis. This is an improvement over Stahl (1999), where no rule learning could be rejected for only 39.7% of the participants.

As another test, we considered just the restriction that $\beta_1 = 0$ while allowing β_0 to be free. The dynamic equation of motion simplifies to

$$w(\rho, t + 1) = \beta_0 w(\rho, t). \quad (2')$$

This hypothesis allows the rule probabilities (φ) to change (when $\beta_0 \neq 1$), but only by deepening ($\beta_0 > 1$) or lessening ($\beta_0 < 1$) the initial log-probability differences across rules, completely independent of the performance information $g(\rho, \Omega^{t+1})$. When we impose this restriction, the aggregate maximized LL decreases to -2151.30 . Twice the difference is distributed χ^2 with 91 degrees of freedom and has a p-value less than 10^{-58} . Thus, we can reject the hypothesis that $\beta_1 = 0$ for all participants. Furthermore, on an individual basis we find that for 46 of 91 (50.55)

¹⁵Since we have $\sigma = 0.25$, this nested Nash model allows some “trembling” to other rules (albeit a negligible amount); the $\bar{\theta}$ parameter is theoretically relevant although of no practical significance.

participants, we can reject the hypothesis that $\beta_1 = 0$. Thus, about half of the participant population appears to have learned which rules were better. The comparable result in Stahl (1999) was only 19%.

If we had only two dimensions or only a small number of rules, then we could easily present a potentially revealing plot of φ over time. However, with five dimensions, it is a challenge to present a picture of how φ changes over time. After an extensive investigation of the modes of the φ distribution, we concluded that the location of the “dominant” mode conveys useful information. Specifically, for each period we identified all grid points for which φ was within 50% of the maximum φ for that period, and computed the φ -weighted average of those grid points; call this $\nu^*(t)$ for period t . Figure 5 displays $\nu_k^*(t)$, $k = 0, \dots, 3$, averaged over the 46 participants with statistically significant β_1 estimates. While aggregation masks individual differences, Fig. 5 sheds light on what rules participants learned. Note that the vertical scale is logarithmic. The initial weight given to the level-2 evidence and the Nash evidence is quite small and increases substantially over time. Indeed, the weight on level-2 evidence increases by as much as 4-fold, while the weight on Nash evidence increases 2.3-fold. This is understandable, since rules with more weight on level-2 evidence yield consistently larger ex post expected utility than other rules, while the aggregated reinforcement of Nash evidence is compromised by the infrequent convergence to Nash equilibrium. These changes are even more dramatic for individuals with above average β_1 estimates. Figure 5 also reveals modest increases in the averaged weights for level-1 and herd

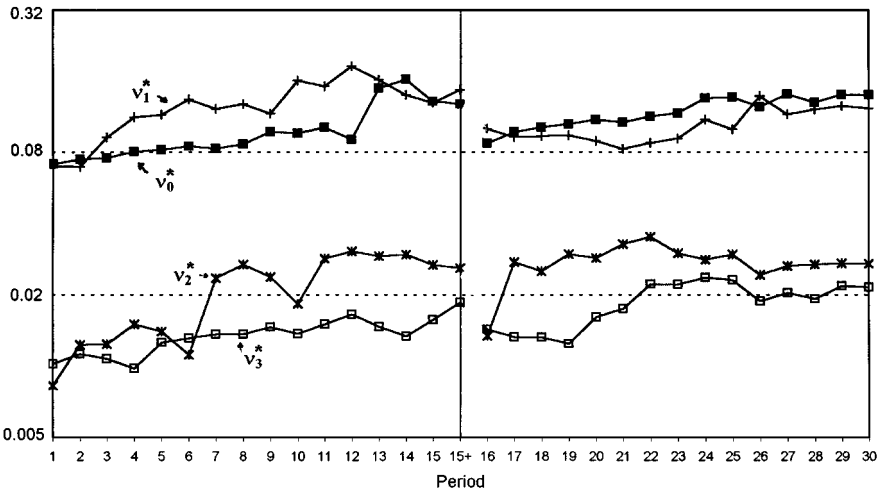


FIG. 5. Coordinates of dominant φ mode for learners.

evidences. It is noteworthy that level-1 evidence is devalued during most of the second run when level-2 evidence is augmented.

RESULT 3. We can reject “Cournot dynamics.”

So-called Cournot dynamics have been popular because of their simplicity and explanatory power (e.g., van Huyck et al., 1994; Cheung and Friedman, 1997, and Friedman et al., 1995). In our context, Cournot dynamics is equivalent to zero weight on all evidence except $y_1(\Omega', \theta)$, and no rule learning.¹⁶ Thus, the reduced model would have only two parameters (ν_1, θ). Maximizing the log-likelihood function with respect to these two parameters for each individual participant and summing over all 91 participants, the aggregated LL decreases to -2829.30 . Compared to the no-rule-learning model, twice the difference is distributed χ^2 with 364 ($= 4 \times 91$) degrees of freedom and has a p-value less than 10^{-67} . Compared to the full model, twice the difference is distributed χ^2 with 637 ($= 7 \times 91$) degrees of freedom and has a p-value less than 10^{-120} . Thus, we can strongly reject the Cournot model in favor of both the no-rule-learning model (but other rules present) and the full rule-learning model. An enhanced Cournot model with learning is also rejected.

RESULT 4. We can strongly reject the hypothesis that the population of participants is homogeneous.

If the population of participants is homogeneous, then there should be a single parameter vector β that applies to all individuals. Estimating the model under this restriction, the maximized LL decreases to -3060.88 . This hypothesis is nested within the full model (entailing 9×90 parameter restrictions), so twice the difference is distributed χ^2 with 810 degrees of freedom, and has a p-value of less than 10^{-146} . Thus, we strongly reject the homogeneity hypotheses. We also tested and rejected the homogeneity hypothesis for each session separately (the smallest p-value is 10^{-22}).

This result indicates that even though the full model contains 819 parameters, we have not overfit the data by allowing too many parameters. There are $91 \times 30 = 2730$ observations, so there remain 1911 degrees of freedom (21 per participant).

RESULT 5. Each of the nine parameters makes a statistically significant contribution in the aggregate.

To test whether, say, $\bar{\nu}_0$ makes a statistically significant contribution in the aggregate, we maximized the log-likelihood function for each participant while holding $\bar{\nu}_0 = 0$, and then summed over all 91 participants. Twice the difference between this restricted LL value and the unrestricted

¹⁶Strict Cournot dynamics would have $\theta = 1$.

value of -1896.44 is distributed χ^2 with 91 degrees of freedom. In this manner, we tested $\bar{\nu}_k$ ($k = 0, \dots, 3$), $\bar{\theta}$, β_1 , and τ . We did not test $\sigma = 0$ due to the singularity of the initial log-propensity function, $w(\cdot, 1)$, at $\sigma = 0$. For tests involving β_0 , see Result 6 below. Table I lists the results. We clearly reject each hypothesis that the corresponding parameter is 0 for all participants. Therefore, we cannot safely drop any of these parameters from the model without a statistically significant loss in explanatory power.

On an individual participant basis, the parameters that are most frequently significantly different from 0 are $\bar{\nu}_1$, $\bar{\theta}$, β_1 , and τ . This suggests that the level-1 evidence receives substantial weight, the recent past is used in forecasting the present, rule performance influences rule choice, and learning is transferred across games.

We hasten to point out that not being able to reject $\bar{\nu}_k = 0$ for some k and some participant does not imply that the corresponding evidence could be dropped from the model for that participant because rule learning could cause ν_k to become significant over time.

b. Individual Parameter Estimates

Table II presents the variance-covariance matrix of the maximum likelihood estimates of the nine parameters over the 91 participants. The simple arithmetic averages (“avg”) and standard deviations (“std”) are given at the bottom of Table II. However, since a logarithmic scale was used for the $\bar{\nu}_k$ parameters, a geometric mean (“geo”) was also computed and reported. These means reveal that the predominant initial evidence

TABLE II
Variance-Covariance of Individual Parameter Estimates

	$\bar{\nu}_0$	$\bar{\nu}_1$	$\bar{\nu}_2$	$\bar{\nu}_3$	$\bar{\theta}$	σ	β_0	β_1	τ
$\bar{\nu}_0$	2.758	0.1174	0.0051	0.0029	-0.0746	-0.0192	0.3747	-1.233	-0.0046
$\bar{\nu}_1$		3.724	0.3301	0.1996	-0.0638	0.6384	-0.1428	-1.405	0.1044
$\bar{\nu}_2$			0.2650	0.1045	-0.0164	-0.0569	-0.0420	-0.1688	0.0518
$\bar{\nu}_3$				0.6874	-0.0537	0.2806	-0.0092	-0.3332	0.0587
$\bar{\theta}$					0.1548	-0.0074	-0.0301	-0.3587	0.0230
σ						2.375	0.0117	0.0589	-0.0061
β_0							0.2136	0.1992	0.0032
β_1								3.573	-0.0562
τ									0.2176
avg:	1.120	1.354	0.1913	0.2580	0.4322	1.145	1.256	1.535	0.4731
std:	1.652	1.919	0.5119	0.8245	0.3912	1.533	0.4596	1.880	0.4647
geo:	0.0793	0.1218	0.0095	0.0089					

weight is on level-1 evidence. The individual parameter estimates for the 91 participants are in the Appendix.

RESULT 6. The distribution of $\bar{\nu}$ weights has three modes.

We sought a parsimonious means of capturing the distribution of individual parameter estimates. In particular, we were curious about the initial distribution of evidence weights $\bar{\nu}$, since they correspond to the SW archetypes. To explore this, we constructed a histogram of the four $\bar{\nu}_k$ estimates using high–low categories¹⁷; thus, our histogram had 16 cells, shown in Table III.

We see that the vast majority (73.6%) of participants placed little initial weight on level-2 and Nash evidence. With 16 cells and 91 observations, under the null hypothesis of a uniform distribution there is a 95% probability that any given cell will contain 10 or fewer observations. Hence, three cells fail this test, meaning that we can reject the hypothesis of a uniform distribution. Indeed, these three cells contain 58 of the 91 participants (63.7%).

It is also insightful to ask whether the parameter estimates are distributed uniformly within each of the three dominant cells, the alternative hypothesis being that they are clustered into modes within the cells. To test this, we compare the mean of the $\bar{\nu}$ estimates in a cell of Table III with the midpoint of that cell and, adjusting for the number of observations, we find that the mean is at least two standard deviations away from the midpoint in the direction of the “low” boundary. Thus, we reject the hypotheses of uniform distributions within cells, suggesting instead that

¹⁷Using the logarithmic scale, $1 + \ln(5/\nu_k)/\ln(4) \geq 4$ defined the “low” category; in other words, the median or higher grid point was declared low. We adjusted these estimates to take account of statistical significance.

TABLE III
Histogram of Adjusted $\bar{\nu}$ Estimates^a

		$(\bar{\nu}_2, \bar{\nu}_3)$				
		LL	LH	HL	HH	
$(\bar{\nu}_0, \bar{\nu}_1)$	LL	15	1	0	2	18
	LH	28	2	4	6	40
	HL	15	0	2	0	17
	HH	9	3	1	3	16
		67	6	7	11	91

Key: L = low; H = high.

^aCoefficients insignificant at 25% jointly tested and restricted to 0.

the means of each cell characterize a definite mode of the overall distribution. These three modes correspond to three "types" of initial log-propensities: (i) "null" ($\bar{v}_k \approx 0$ for all k), (ii) level-1 ($\bar{v}_1 > 0$ and $\bar{v}_k \approx 0$ for all $k \neq 1$), and (iii) herd followers ($\bar{v}_0 > 0$ and $\bar{v}_k \approx 0$ for all $k \neq 0$).

An alternative to this crude histogram would be kernel density estimation. However, only 91 noisy observations and four dimensions makes kernel density estimation unreliable. On the other hand, for just the two dimensions (\bar{v}_0, \bar{v}_1), the 91 observations are sufficient for reliable kernel estimation, and we find strong statistical evidence for the three modes close to the mean of the data in the three dominant cells of our crude histogram. Figure 6 presents the estimated density of \bar{v}_0 and \bar{v}_1 with three clear modes.

Given these three modes, it is natural to look for correlations between these modes and the estimates of the other parameters. The mean estimates of $\bar{\theta}$, σ , β_0 , and τ for these three modes are statistically indistinguishable. However, we found the learning parameter β_1 to be signifi-

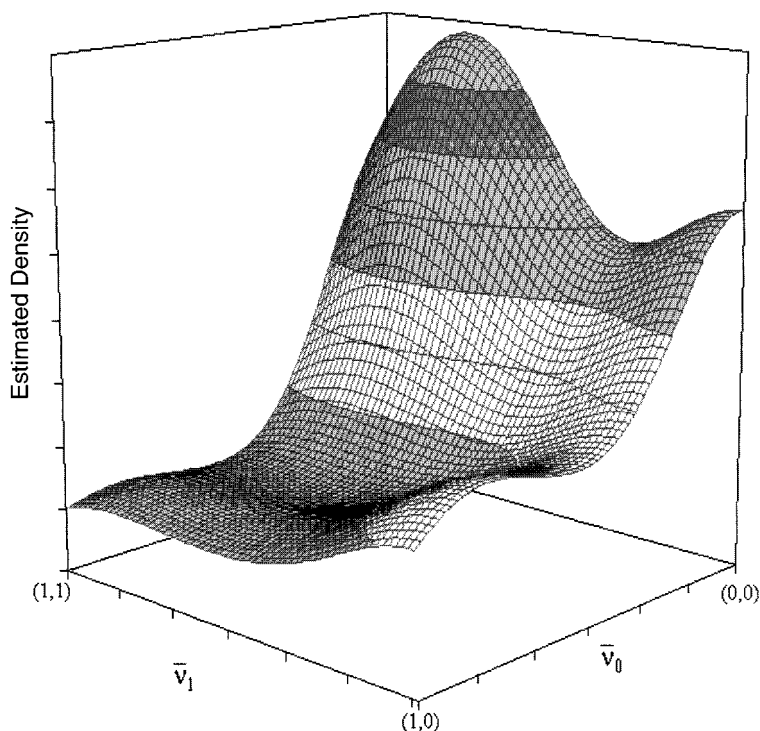


FIG. 6. Kernel estimated density of (\bar{v}_0, \bar{v}_1) .

cantly larger for the first and third mode. Since it turns out that level-1 evidence is reinforced, it is important that those participants who do not initially put much weight on level-1 evidence eventually learn to do so; hence, it makes sense that β_1 is higher for the first and third mode. We also found that the mean estimates over the 13 other cells of our histogram were significantly higher for σ and lower for β_1 and τ . This finding suggests that the 58 participants in the three dominant modes had much sharper initial dispositions (low σ), tended to respond more to performance (high β_1), and transferred more of what they learned across the two runs than the other participants.

RESULT 7. We cannot reject asymptotic stability for most participants.

Asymptotic stability requires $\beta_0 < 1$. For 60 (66%) of the participants, the estimated β_0 exceeds 1, and the mean also exceeds 1, which seems to suggest explosive dynamics for the majority of participants.¹⁸ We therefore reestimated the model with the restriction that $\beta_0 \leq 1$, and found that we could not reject the restriction at the 5% level for 69 (75.8%) of the participants; that is, only one-fourth of the participants appear to have explosive dynamics. On the other hand, the aggregate test for all participants rejects the restriction with a p-value of 2×10^{-14} .

5. DISCUSSION

This paper tested a theory of boundedly rational behavioral rules, and a theory of *rule learning* based on past performance. The boundedly rational rules can be interpreted as weighing *evidence* for and against available actions based on archetypal models of the other players. An econometric model was specified and an experiment was designed to fit and test this model.

Our model fits the data much better than random noise or an error-prone Nash model. We strongly reject the hypothesis of no rule learning and the hypothesis of Cournot dynamics in favor of our model. The maximum likelihood estimates of the parameters reveal substantial heterogeneity in the population of participants both in the initial log-propensities over rules and the extent of learning. A statistical test of homogeneity strongly rejects that hypothesis. After exhaustive testing of each of the nine parameters of

¹⁸To ensure convergence of the β_0 estimates, an upper bound of 2.0 was imposed on the range. For five participants, the estimated β_0 hits this boundary, but without the upper bound, the optimization subroutine tended to increase β_0 indefinitely; however, the log-likelihood values increased only negligibly for $\beta_0 > 2$, and the other parameter estimates changed negligibly. Hence, the upper bound is innocuous.

the rule-learning model, we rejected the hypotheses that any one of them could be dropped from the model.

Herd evidence y_0 and level-1 evidence y_1 receive the most initial weight on average. However, one should not conclude that level-2 evidence (y_2) and Nash evidence (y_3) are unimportant, since they can become important via learning. The average estimate of β_1 is 1.53, so substantial learning can occur. Further, examining the reinforcement function $g(\nu, \theta, \Omega^{t+1})$, we found that the “99% best” rules in our (ν, θ) rule space generally included high values for ν_2 and ν_3 . Knowing that the no-rule-learning hypothesis was strongly rejected, there is no doubt that this reinforcement of level-2 and Nash evidence contributed to the success of the model. We also presented graphical evidence (Fig. 5) showing substantial changes in the weight given level-2 and Nash evidence over time, consistent with rule learning and increasing sophistication.

Using a histogram approach and kernel density estimation, we were able to identify three modes in the distribution of the initial evidence weights $\bar{\nu}$. This suggests that it might be possible to specify a parsimonious finite mixture model that explains the data as well (when accounting for the difference in the number of parameters).

While we could not reject asymptotic stability ($\beta_0 \leq 1$) for three-fourths of the participants, it may trouble some readers that 25% appear to have explosive dynamics. However, it should be noted that the only way for the probabilistic choice function to eventually put unit mass on one rule (such as the Nash rule) is to have explosive dynamics.

An apparent weakness of our learning theory is the implicit assumption that participants evaluate the whole five-dimensional space of rules R and update their rule probabilities φ accordingly. It might seem more realistic to assume that the players gather incomplete samples of rule performance measures depending on similarity or closeness with the rules recently used. Unfortunately, we cannot directly observe this sampling and evaluation process. On the other hand, rules that are “close” in the brains of human subjects are not necessarily those that are close in our parametric representation (and vice versa). In other words, a rule that involves a large change in the parameter space of our representation is not necessarily more distant, and hence less likely to be evaluated. Stinchcombe (1997) has shown that artificial neural networks are universal approximators even for arbitrarily small parameter sets; that is, small changes in the parameter weights of a neural network can span a large function space. Therefore, it is theoretically possible that local experimentation in the weight-space of a brain’s neural network could in fact span a space of rules as large as and similar to our five-dimensional space of evidence-based rules.

APPENDIX: INDIVIDUAL PARAMETER ESTIMATES

ID	\bar{p}^0	\bar{p}^1	\bar{p}^2	\bar{p}^3	$\bar{\theta}$	σ	β_0	β_1	τ	LL
1	0.5408 ^o	0.0000	0.0000	0.0000	1.0000*	0.2500	0.8726	0.0000	0.9998	-30.748
2	2.5038	0.1977 [†]	0.0636*	0.0000	0.0000	0.2500	1.2112	0.5988 [†]	0.0868 [†]	-23.110
3	0.0000	4.4411 [†]	2.5065 [†]	0.9818 [†]	0.0480	0.5972	1.6144 ^o	0.2428 [†]	0.6661*	-10.737
4	2.4725	4.9986*	0.9942*	1.4766	0.0000	0.3547	0.8123	0.1248 ^o	0.0221 ^o	-15.233
5	0.0625 [†]	4.9971 [†]	0.5292 [†]	0.4845*	0.9968 [†]	0.2500	1.1034	0.6556*	0.0000	-0.227
6	0.0000	1.4120 [†]	0.0200 ^o	0.0000	0.7680 [†]	0.4754	2.0000	5.0000*	1.0000 [†]	-7.235
7	5.0000	5.0000*	1.2286	1.5173 ^o	0.1148*	5.0000	2.0000 [†]	0.1747	0.0679*	-5.668
8	2.4729	2.4368 [†]	0.9038 ^o	0.2573 ^o	0.0000	0.2500	1.1127	0.4064 [†]	1.0000 [†]	-9.389
9	0.0008	4.9945 [†]	0.0000	0.0131	0.0000	0.2506	0.4292	1.0409 [†]	0.0002	-9.842
10	0.1557 [†]	5.0000*	0.0023	0.9507*	1.0000 [†]	5.0000	1.9056	0.0041	0.0519*	-10.549
11	1.6226	0.0056	0.0010	0.0000	0.9364	4.2897	1.9745	0.0275 ^o	1.0000*	-20.167
12	2.3755	0.2457 [†]	0.0012	0.0515	0.0000	0.2533	0.8049	0.1027*	0.9996*	-17.351
13	0.0000	1.8881	0.0276	0.0041	0.1390 [†]	3.6635	1.6504	0.0362	0.9998*	-15.107
14	0.0000	0.0836 [†]	0.0026*	0.0017	0.8155 [†]	0.2500	1.1202 [†]	5.0000 [†]	0.1685 [†]	-19.759
15	2.4961	0.2822*	0.0627*	0.0000	0.0000	0.2500	0.9333	0.8335 [†]	0.9184 [†]	-9.118
16	0.3372	0.2125	0.0644	0.0051 ^o	0.6182	2.3719	2.0000	0.0000	0.0436 ^o	-20.501
17	0.1779	2.3597 [†]	0.0010	0.0098 [†]	0.7380*	1.6222	1.9991*	0.3750 ^o	0.0000	-6.835
18	0.9154 ^o	0.0194	0.0008	0.0020	0.0160	0.3176	1.1566 ^o	4.9996 [†]	0.0000	-13.717
19	4.9978*	0.0000	0.0000	0.0123	0.8029*	0.5256	0.8164	0.0000	0.0000	-18.858
20	0.0000	4.8982 [†]	0.0754	0.0359	0.0444	0.2501	1.2870	2.2908 [†]	0.4383 [†]	-0.000
21	4.0110	0.4785	0.0068	0.0140	0.1299	2.4506	1.9996	0.2418	0.0033 ^o	-7.461
22	0.0000	0.0770*	0.0000	0.0228*	0.2410	0.2932	1.3692	4.9900*	0.0036*	-14.091
23	0.0000	4.9990*	2.1739 [†]	0.0088	1.0000	0.2501	0.4507	0.0575	0.0336*	-22.408
24	0.0000	0.0000	0.0593*	0.0001	0.0000	0.7719	0.9035	0.0577	1.0000*	-32.326
25	0.0096	0.0000	0.0000	0.0014	0.0447 [†]	0.2514	1.2981 [†]	3.6047*	0.8085*	-43.799
26	0.0040*	0.0000	0.0064 ^o	0.0000	0.9127 [†]	0.6952	1.3744 [†]	1.6411 [†]	0.9350*	-26.881
27	0.3714	5.0000 [†]	0.0000	0.0000	0.5993	2.5876	1.0722	0.0000	1.0000	-24.947
28	0.6240	0.1001 [†]	0.0000	0.0103	0.0000	0.2500	1.1845 ^o	2.1288*	0.0153 ^o	-31.331
29	0.0087	0.0000	0.0012	0.0000	1.0000 [†]	0.2500	0.8493	0.6704*	0.9987 [†]	-24.132
30	0.0000	4.9997 [†]	0.0000	0.0012	0.2720	0.2500	0.3618	0.4897*	1.0000*	-15.816
31	0.1545	0.0000	0.1242 [†]	4.0937 [†]	0.0387	5.0000	1.5068 [†]	0.0099	1.0000 [†]	-15.783
32	0.0118	0.0000	0.0000	0.0037	0.8475	0.2999	1.4574	2.8662	1.0000	-44.614
33	0.0001	5.0000 [†]	0.0001	0.0078	0.0000	0.2500	0.1498	0.1621	1.0000 [†]	-28.952
34	5.0000 [†]	1.6528 [†]	0.0384	0.1075	1.0000 [†]	0.2500	0.9181	0.0000	1.0000	-5.694
35	0.0442	0.0039*	0.0012	0.0000	0.0010	4.9990	1.3331*	0.0095	1.0000*	-42.389
36	5.0000	0.6763 [†]	0.1394*	0.0000	0.3720	0.2500	0.6182	0.0107	0.0000	-25.535
37	0.0021	0.3760 [†]	0.0000	0.0000	0.4976 [†]	0.2500	1.0286	2.6208	1.0000	-18.309
38	0.0721 ^o	0.0102 [†]	0.0000	0.0000	0.8219 [†]	0.2500	1.1850	3.0163 [†]	0.7813 [†]	-31.967
39	0.0000	0.0025	0.0000	0.0000	1.0000	0.2507	1.3836	4.9836	0.0000	-44.853
40	0.0000	0.9091*	0.0204	0.0021	0.1821 [†]	1.0786	1.5712	0.9160	0.1106 [†]	-21.850
41	0.0000	5.0000 [†]	1.7114 ^o	0.0000	1.0000 [†]	0.2500	0.5904	0.8460	0.2451 [†]	-11.346
42	0.0000	4.9990 [†]	0.0000	0.0000	0.0511	2.4469	1.1945*	0.0066	0.2063*	-24.424
43	0.5839	0.0216*	2.5775 [†]	0.0012	0.1535*	0.2502	0.9575	2.3487 [†]	0.0000	-4.742
44	0.0000	0.5945 ^o	0.0862	0.0825	0.4837	4.1516	2.0000 ^o	0.2990	1.0000 ^o	-17.525
45	1.9268*	0.0000	0.0128 ^o	0.0172	0.3699*	0.6839	1.8660 [†]	4.9990	0.7342	-5.410

ID	\bar{v}^0	\bar{v}^1	\bar{v}^2	\bar{v}^3	$\bar{\theta}$	σ	β_0	β_1	τ	LL
46	2.7593 [†]	0.0634 ^o	0.0000	0.0000	0.0642	0.3885	1.9299	4.9989*	0.9068	-22.125
47	0.0101	0.3426 [†]	0.0000	0.0000	0.0952	0.2501	0.9071	0.0000	0.0013	-29.268
48	2.0026*	0.0002	0.1977*	0.0000	0.3901*	4.4643	1.5719	0.0036 ^o	0.1303*	-32.949
49	0.1562 [†]	4.9975 [†]	0.0028	0.0097	0.5970	4.6817	1.3992*	0.0791*	0.9004 [†]	-16.127
50	4.9992	5.0000 [†]	0.0543	0.0000	0.2650	2.5071	1.1539	0.0000	0.0343	-29.060
51	0.0152 [†]	0.0009	0.0043 ^o	0.0012*	1.0000 [†]	0.2500	0.7468	0.7763 [†]	0.0000*	-27.068
52	0.1102*	0.0022	0.0026	0.0027	0.7794*	0.2523	0.9677	1.3121	0.6872 [†]	-26.297
53	0.4485 [†]	0.0100*	0.0000	0.0025	1.0000 [†]	0.3440	1.9834 [†]	4.9999 [†]	0.9925	-31.425
54	0.0000	0.2176 [†]	0.0000	0.0000	0.1795	0.2500	0.7955	0.0066	0.0000	-35.825
55	3.4106*	0.0000	0.0000	0.0069	0.4500*	0.3325	1.3563*	4.9997*	1.0000 [†]	-18.367
56	0.1524*	0.1368 [†]	0.0218*	0.0023	0.0310	3.4046	2.0000*	0.0134*	0.0000	-37.722
57	0.7113	1.2696 [†]	0.0027 ^o	0.0249*	0.2661 [†]	0.5056	1.7863*	4.8156 [†]	0.0717*	-23.277
58	0.0006	0.0012	1.2173 [†]	5.0000	0.0000	0.2500	0.6789	0.4170 [†]	0.0015 ^o	-6.561
59	0.0000	0.1217 [†]	0.0101*	0.0161*	0.7319 [†]	0.2602	1.0671	4.9999 [†]	1.0000*	-15.697
60	2.7408	1.5239	0.9900 ^o	0.0118	0.9666*	0.2500	0.8777	0.6148 [†]	0.0069	-4.061
61	0.3118	5.0000 [†]	0.1443	0.0010	0.0312	0.3302	0.9002	0.0182	0.9952*	-23.156
62	0.0000	0.0018	0.0000	0.1046 [†]	1.0000 [†]	0.6843	0.8165	0.9360 [†]	0.0000	-16.074
63	1.7728*	0.0091	0.0000	0.0000	0.0003	0.2999	1.5141	2.5531*	0.0265	-42.258
64	0.0825	1.4993*	0.0012	0.0000	0.8662 [†]	4.9990	1.9990	0.0402	1.0000*	-20.602
65	0.2375 [†]	0.0314 [†]	0.0000	0.0000	1.0000	0.2500	0.7624	0.8180 [†]	0.0000	-24.642
66	3.8120 [†]	4.5427 [†]	0.0024	1.1055 [†]	0.0000	3.1310	1.7161 [†]	0.0204 [†]	0.2020*	-21.404
67	0.0000	1.3477 [†]	0.0124 [†]	0.0348 ^o	0.9894 [†]	0.3027	1.7306*	3.9892*	0.9994 [†]	-21.377
68	0.0000	0.2974*	0.0660*	0.0146	0.0437	3.8471	1.9989	0.0041	0.0000	-31.950
69	4.9990 [†]	0.0000	0.0591*	0.0012	0.5483	0.2501	0.6415	1.1189 [†]	1.0000 [†]	-12.059
70	3.6122*	1.8457 [†]	0.2781 [†]	0.0704 ^o	0.9999 [†]	0.2501	1.1192	0.3478 ^o	0.9399	-2.743
71	0.0097	0.0000	0.0000	0.0000	0.2037*	0.2879	1.7774 ^o	4.9893*	0.0000	-43.732
72	5.0000	0.1570 [†]	0.0000	0.0000	0.0000	0.2500	0.8264	0.1013	1.0000	-25.207
73	0.0000	0.0182	0.0014	0.0379	0.7706 [†]	0.2558	1.6078 [†]	4.9696 ^o	0.0000	-37.219
74	0.1570 [†]	0.0000	0.0006	0.0000	0.1901 [†]	0.3219	1.6613 [†]	3.3650 [†]	1.0000 [†]	-28.528
75	2.5040 [†]	4.0819 [†]	0.5188 [†]	1.6948 [†]	0.0000	0.2500	0.5901	0.5625 [†]	0.0000	-11.477
76	0.0348	0.0375 [†]	0.0000	0.0119	1.0000*	0.2500	0.8631	0.6736 [†]	0.0339*	-23.610
77	1.7452*	0.0937	0.0119	0.0146	1.0000*	0.2500	1.9999	0.0000	0.0000	-19.141
78	0.3290	2.7803*	0.1246	0.0809	0.1616	0.9194	1.1698	0.0010	0.0000	-15.864
79	0.0000	0.6532 [†]	0.0134	0.0000	0.1962	0.8838	0.9413	0.0000	1.0000	-32.711
80	0.9836	0.2932 [†]	0.0028	0.0000	0.0148	0.2500	1.1703	3.4259 [†]	0.0000	-7.412
81	0.0000	0.1251 [†]	0.0012	0.0352 [†]	0.5036 [†]	0.2500	1.2084	3.0037 [†]	0.0000	-17.391
82	0.5162 [†]	0.0000	0.0000	0.0158*	0.5516	0.2501	1.3538 [†]	4.9990 [†]	1.0000	-14.026
83	4.9540 [†]	0.0813	0.0000	0.0000	0.0967 [†]	0.2500	1.1017	4.9999	0.03252	-8.780
84	3.0285	4.9805 [†]	0.0132 ^o	0.9418 [†]	0.7177 [†]	2.7134	1.5120 [†]	0.0518*	0.0002	-16.427
85	0.0000	0.1841 [†]	0.0000	0.0000	0.0910	0.7047	0.9061	0.0000	1.0000	-39.521
86	0.1443	0.0000	0.0000	0.0000	0.1744 [†]	0.2501	1.3706	4.4997 [†]	0.9974*	-25.421
87	4.9997 [†]	0.0597	0.0021	0.0000	0.0000	0.2500	1.3772	3.0778	1.0000	-17.958
88	0.3047	0.1069 [†]	0.0039	0.0000	0.7543 [†]	0.2500	0.8871	0.6666 [†]	1.0000*	-22.962
89	0.0000	0.0787 [†]	0.0000	0.0415*	0.1431 [†]	0.2500	1.1354	3.3992*	0.0000	-28.235
90	0.5741	1.6877 [†]	0.2061 [†]	0.2124 ^o	1.0000 [†]	5.0000	1.7180 [†]	0.0027 ^o	0.0898*	-15.789
91	0.3177	1.1829	0.0003	3.7984 [†]	0.5434 [†]	2.2174	1.2515*	0.0868	0.0008	-22.203

^oSignificant at 10% level; *significant at 5% level; [†]significant at 1% level.

$\beta_0 \leq 1$; σ was not tested.

REFERENCES

- Anderson, S., Goeree, J., and Holt, C. (1997). "Stochastic Game Theory: Adjustment to Equilibrium With Bounded Rationality," mimeo, Department of Economics, University of Virginia.
- Anderson, S., Goeree, J., and Holt, C. (1999). "Bounded Rationality in Markets and Complex Systems," *Experiment. Econom.*, forthcoming.
- Bernheim, B. D. (1984). "Rationalizable Strategic Behavior," *Econometrica* **52**, 1007–1028.
- Binmore, K. (1987). "Modeling Rational Players, I and II," *Economics and Philosophy* **3**, 179–214; **4**, 9–55.
- Camerer, C., and Ho, T. (1997). "EWA Learning in Games: Preliminary Estimates from Weak-Link Games," in *Games and Human Behavior: Essays in Honor of Amnon Rapoport* (R. Hogarth, Ed.).
- Camerer, C., and Ho, T. (1999). "Experience-Weighted Attraction Learning in Normal-Form Games," *Econometrica* **67**, 827–874.
- Cheung, Y-W., and Friedman, D. (1997). "Individual Learning in Games: Some Laboratory Results," *Games Econom. Behav.* **19**, 46–76.
- Cooper, D., and Feltovich, N. (1996). "Reinforcement-Based Learning vs. Bayesian Learning: Comparison," mimeo, University of Pittsburgh.
- Crawford, V. (1994). "Adaptive Dynamics in Coordination Games," *Econometrica* **63**, 103–143.
- Erev, I., and Roth, A. (1998). "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria," *Amer. Econom. Rev.* **88**, 848–881.
- Friedman, D., Massaro, D., and Cohen, M. (1995). "A Comparison of Learning Models," *J. Math. Psychol.* **39**, 164–178.
- Hofbauer, J., and Sigmund, K. (1988). *The Theory of Evolution and Dynamical Systems*, Cambridge, UK: Cambridge University Press.
- Jordan, J. (1991). "Bayesian Learning in Normal Form Games," *Games Econom. Behav.* **3**, 60–81.
- Kalai, E., and Lehrer, E. (1993). "Rational Learning Leads to Nash Equilibrium," *Econometrica* **61**, 1019–1045.
- Machina, M., and Schmeidler, D. (1992). "A More Robust Definition of Subjective Probability," *Econometrica* **60**, 745–780.
- McFadden, D. (1974). "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics* (P. Zaremba, Ed.). New York: Academic Press.
- McKelvey, R., and Palfrey, T. (1995). "Quantal Response Equilibria for Normal Form Games," *Games Econom. Behav.* **10**, 6–38.
- McKelvey, R., and Palfrey, T. (1998). "Quantal Response Equilibria for Extensive Form Games," *Experiment. Econom.* **1**, 9–42.
- Mookherjee, D., and Sopher, B. (1994). "Learning Behavior in an Experimental Matching Pennies Game," *Games Econom. Behav.* **7**, 62–91.
- Mookherjee, D., and Sopher, B. (1997). "Learning and Decision Costs in Experimental Constant Sum Games," *Games Econom. Behav.* **19**, 97–132.
- Nagel, R. (1995). "Unraveling in Guessing Games: An Experimental Study," *Amer. Econom. Rev.* **85**, 1313–1326.

- Pearce, D. (1984). "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica* **52**, 1029–1050.
- Rapoport, A., Erev, I., Abraham, E., and Olson, D. (1997). "Randomization and Adaptive Learning in a Simplified Poker Game," *Organizat. Behav. Human Decision Process.* **69**, 31–49.
- Roth, A., and Erev, I. (1995). "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games Econom. Behav.* **8**, 164–212.
- Sonsino, D., Erev, I., Gilat, S., and Shabtai, G. (1998). "On the Likelihood of Repeated Zero-sum Betting by Adaptive Human Agents," Working Paper, Faculty of Industrial Engineering and Management, Technion, Israel.
- Stahl, D. (1993). "The Evolution of Smart_n Players," *Games Econom. Behav.* **5**, 604–617.
- Stahl, D. (1996). "Boundedly Rational Rule Learning in a Guessing Game," *Games Econom. Behav.* **16**, 303–330.
- Stahl, D. (1999). "Evidence Based Rules and Learning in Symmetric Normal-Form Games," *Int. J. Game Theory* **28**, 111–130.
- Stahl, D., and Wilson, P. (1994). "Experimental Evidence of Players' Models of Other Players," *J. Econom. Behav. Organiz.* **25**, 309–327.
- Stahl, D., and Wilson, P. (1995). "On Players Models of Other Players: Theory and Experimental Evidence," *Games Econom. Behav.* **10**, 218–254.
- Stinchcombe, M. (1997). "Neural Network Approximation of Continuous Functionals and Continuous Functions on Compactifications," mimeo, Department of Economics, University of Texas.
- Thorndike, E. (1898). "Animal Intelligence: An Experimental Study of the Associative Processes in Animals," *Psycholog. Monogr.* **2**.
- Van Huyck, J., Cook, J., and Battalio, R. (1994). "Selection Dynamics, Asymptotic Stability, and Adaptive Behavior," *J. Polit. Economy* **102**, 975–1005.