Naive Bayesian Learning and Adjustment to Equilibrium in Signaling Games

Jordi Brandts and Charles A. Holt*

ABSTRACT

This paper uses laboratory and simulation techniques to examine behavior in two signaling games with different payoff structures. Each game has two sequential equilibria in pure strategies, one of which is ruled out by standard "refinements." The behavior of human subjects approaches the more refined equilibrium in one of the games, but it approaches the less refined equilibrium in the other game. This difference in subjects' decisions is predicted by a simple Bayesian learning process. The period-by-period pattern of adjustment is tracked by computer simulations that incorporate Bayesian learning, logistic decision errors, and some strategic anticipation.

Journal of Economic Literature classification: C72, C92
keywords: experiments, game theory, learning, simulations

1. INTRODUCTION

The presence of multiple Nash equilibria in all but the simplest games has produced a quest for a Holy Grail that would indicate a unique, plausible outcome in all games. The most common approach has been to develop theoretical criteria for eliminating those equilibria that are based on "unreasonable" beliefs. The theory was developed through a series of well known examples of games in which successively stronger restrictions on beliefs were used to rule out unwanted equilibria. This introspective approach selected certain equilibria in these hypothetical examples, and these equilibria essentially played the role of data to be accommodated by the

---

* Instituto de Análisis Económico (CSIC) in Barcelona and University of Virginia, respectively. This paper was profoundly influenced by the work of Richard Cyert and Morris DeGroot on Bayesian learning and adaptive modeling in economics. We acknowledge financial support from the Spanish DGCICYT (PB92-0593 and PB93-0679, the U.S. NSF (SBR-9320044), and the CIRIT, Generalitat de Catalunya. We acknowledge helpful comments and criticisms from José Canals, Russ Cooper, Richard Gilbert, Richard McKelvey, Jennifer Novak, Tom Palfrey, Isabel Sánchez, and Bill Zame. All correspondence should be sent to Charles A. Holt, Department of Economics, Rouss Hall, University of Virginia, Charlottesville, Virginia 22901, USA.

theory. Experiments with human subjects can be used to generate data based on actual decisions, which can guide further theoretical work.

The approach taken in the theoretical literature involves considering a particular equilibrium and then determining whether the beliefs that sustain it are reasonable.[1] This "refinement" logic starts from the perspective of a specific equilibrium and considers the reasonable beliefs that a player may have after seeing a deviation.[2] The equilibrium is not rejected if reasonable beliefs motivate the responses to any deviation that deter it. Although theoretical arguments can be useful in ruling out unreasonable beliefs, the increasingly complex refinements have become less convincing in the absence of direct empirical observation.

Our approach to equilibrium selection is to focus on beliefs that develop during adjustment to an equilibrium rather than on beliefs that are reasonable once an equilibrium is reached. Brandts and Holt (1992, 1993) report data from experiments with signaling games in which subjects gain experience with a series of different partners. The experience gained during the adjustment to equilibrium seems to determine the equilibrium selected. In some of these games, the effect of experience is to reinforce beliefs that are ruled out by the refinement approach, and consequently, behavior converges to the *less* refined equilibrium. In this paper we use simulation techniques to examine the process of learning and adjustment in two of the games first reported in Brandts and Holt (1993) in an effort to clarify the behavioral explanation of why some refinements are not good predictors.

The way in which subjects gain experience in a series of matchings with different partners in the laboratory is similar to the strategic interaction modeled in the theoretical literature on evolutionary behavior and learning dynamics. The standard assumption in this literature is that players do not attempt to influence other players' future actions. Naive learning in this context can result in behavior that converges to a Nash equilibrium (Kalai and Lehrer, 1993, and Jordan, 1991). Friedman (1991), Mailath (1992), and Marimon and McGrattan (1995) survey theoretical work in this area.

---

[1] For a discussion of this approach, see Kohlberg and Mertens (1986) and Cho and Kreps (1987).

[2] A theoretical criterion for systematically ruling out a subset of Nash equilibria is known in the literature as a refinement.

There are many ways to model learning in experimental games. Perhaps the simplest assumption is that players learn to use decisions that have yielded high payoffs in the past. This approach requires minimal rationality; agents simply follow the "path of pleasure" as successful strategies gain strength in an evolutionary process. Friedman (1992) reports results from normal-form game experiments designed to test some elements of evolutionary game theory. Crawford (1991) provides an evolutionary interpretation of coordination game experiments. Roth and Erev (1995) report computer simulations in which agents choose a particular decision with a probability that equals the ratio of the sum of payoffs previously obtained with this decision to the sum of payoffs previously obtained from *all* feasible decisions. These models often do not require that players know anything about other's payoffs.

Instead of assuming that players simply tend to repeat decisions that have yielded high payoffs in the past, we take the more structural approach of modeling players who update beliefs about others' decisions, and then calculate expected payoffs for each of their own decisions.[3] There are different ways of modeling this updating process. Cooper, Garvin, and Kagel (1994), for example, use an adaptive rule.[4] We also model the formation of beliefs, but we use a Bayesian rule. Since only one of the other player's decisions is observed each period, the observation can be thought of as the realization of a multinomial variable with probabilities associated with each decision. The natural conjugate prior for this distribution is the Dirichlet (multivariate Beta).[5] This Bayesian approach is naive in the sense that subjects are assumed to make decisions that maximize expected payoffs given these beliefs derived from past decisions, without strategically anticipating future decisions.

In this paper we first examine whether simulations based on the naive Bayesian approach by itself are able to track our experimental data in a satisfactory way. We then explore whether

---

[3] We were influenced by Cyert and DeGroot (1987), which contains a number of important applications of Bayesian learning in economic models.

[4] Other experimental work that explicitly studies learning is relatively recent, but plentiful. Selten and Stoecker (1986) use a learning-theoretic approach to analyze experiments with finite repetitions of prisoner's dilemma games. For work in progress, see Ball and Gardner (1993), Holt (1993), McCabe and Mukherji (1993), Rassenti, Reynolds, Smith and Szidorovszky (1993), Sopher and Mookherjee (1993), and Van Huyck and Battalio (1993).

[5] See DeGroot (1970), section 9.8.

the addition of some strategic anticipation to the naive Bayesian model improves the tracking of the human data. Since subjects make mistakes even in simpler non-strategic situations, we introduce decision errors and analyze their effect on the simulations. These errors are modeled as logistic.[6]

Signaling games constitute a simple but highly interesting way of representing situations with asymmetric information. They have become a standard tool of analysis, comparable to the battle-of-the-sexes game and the prisoner's dilemmma, and are often used in applied theoretical work to model a variety of different economic situations. Examples of this kind of work are Spence's (1973) job market model and Myers and Majluf's (1984) analysis of corporate investment and capital structure. This applied use of signaling models makes is an important motivation for analyzing them through laboratory methods. We believe that the kind of learning model that we use in this paper can be applied usefully to study behavior in a variety of games. Succesful applications of related simulation models to the study of other experimental games can be found in Cooper, Garvin and Kagel (1994), Gale, Binmore and Samuelson (1995) and Roth and Erev (1995). In games in which subjects do not only focus on their own payoffs, like public goods and bargaining games, features not related to learning may have to be incorporated to yield a satisfactory explanation of behavior.

Section 2 describes an experiment that provides support for the intuitive criterion and stronger refinements. Section 3 contains an examination of the time pattern of adjustment, which suggests that the refinement may be working for the wrong reason. Both the initial pattern of out-of-equilibrium decisions and the convergence toward the intuitive equilibrium are consistent with a simple adaptive explanation of behavior. This explanation, which does not rely on the refinement logic, is made precise in a naive Bayesian simulation model, also described in section 3. The following section analyzes a different game, with the same configuration of pure-strategy pooling equilibria. For this game, both human and simulated behavior did not tend towards the refined equilibrium. Section 5 contains a more detailed discussion of the relationship

[6] Anderson (1993) used data from a sequential-choice experiment to estimate a model of decision making with logistic errors.

between human decisions and our computer simulations. The final section contains the conclusion.

2. Game N: The Refinement Logic and Experimental Results

The signaling games to be discussed have two stages: one player sends a signal and the other makes a response. The first player, who will be called the proponent, has a preference type (A or B) that is randomly determined. Each type is equally likely *ex ante*, and the proponent finds out his or her type before choosing a signal: I or S. The respondent sees the signal *but not the proponent's type*, and then makes a response: C, D, or E. The payoffs depend on the proponent's type, the signal, and the respondent's decision.

Table 1. Game N
(proponent's payoff, respondent's payoff)

	response				response		
	C	D	E		C	D	E
type A sends I	45, 30	15, 0	30, 15	type A sends S	30, 90	0, 15	45, 15
type B sends I	30, 30	0, 45	30, 15	type B sends S	45, 0	15, 30	30, 15

Consider the payoffs for Game N that are shown in table 1, with the payoff for the proponent listed first in the payoff pair for each outcome. The top row is for the case of a type A proponent, and the bottom row is for the case of a type B proponent. The proponent either sends an I signal, with payoffs given on the left side of the table then determined by the response column: C, D, or E. Similarly, the response to an S signal determines an outcome with payoffs given on the right.

There are two equilibria in this game, both of which involve pooling. When a proponent sends signal I regardless of his or her type, the respondent should assign a .5 probability to each proponent type. In this case, the best response to I is C, since the respondent's payoff of 30 in the C column on the left side of table 1 dominates response E and is higher than the "fifty-fifty" mixture of 0 and 45 that results from response D. In this equilibrium, the type A proponent would earn 45, and the type B proponent would earn 30. The respondent's equilibrium strategy also specifies a D response to an S signal. This D response, which is optimal if the deviant is thought to be a type B, will prevent either proponent type from deviating.

The intuition behind the (I,C) outcome is more easily understood in terms of the "beer-quiche" story that motivated theoretical work on similar games without the E response. The proponent is strong (A) or weak (B), and knowing this, decides whether to eat quiche (S) or drink beer (I) for breakfast. The respondent sees the breakfast but not the type, and decides whether to concede (C) or duel (D). The payoffs imply that the respondent would rather concede to the strong type and duel the weak type, regardless of the signal. In the I equilibrium, both proponent types drink beer. Neither type would deviate because they expect the duel (D) punishment that is optimal if the respondent interprets quiche eating as a sign of weakness.

There is a second pooling equilibrium for game N, with both types of proponent sending S. The S signal is answered with C, which yields payoffs of 30 for the type A proponent and 45 for the type B proponent. This outcome is supported by a D response to the I signal. The D response is best if the respondent believes that the deviation was very likely to have come from a type B player. Although this equilibrium is sequential, Cho and Kreps (1987) argue that these beliefs are unreasonable in the sense that the type B player earns 45 in the (S,C) equilibrium outcome, and a deviation to message I can never yield a payoff above 30. Hence a deviant I signal should be thought to have come from a type A proponent, which calls for a C response, which in turn gives the type A-proponent an incentive to deviate from the equilibrium. Therefore, the equilibrium involving the S messages is sequential but not "intuitive", as defined by Cho and Kreps.[7]

The argument just given cannot break the other (I,C) pooling outcome, since it is supported by beliefs that the deviant to S is a type B proponent, who could possibly obtain a payoff of 45 from deviating that is greater than the 30 obtained in the (I,C) equilibrium. Therefore, the equilibrium involving the I signals is both sequential and intuitive.

---

[7] Cho and Kreps (1987) show that intuitive equilibria are a subset of sequential equilibria for these signaling games.

We used game N in an experiment with two groups of twelve subjects (a thirteenth subject was used as a monitor to throw dice and verify our procedures). Six of the subjects were proponents, with their type determined by a random device. After learning their own type for the matching, each proponent selected a signal, which was communicated to the matched respondent. Finally, the response was sent back to the matched proponent, and the proponent's type was communicated to the respondent, so that each person could calculate his or her payoff from the table for game N. In order to induce risk neutrality, the payoffs listed for game N were in "points" or lottery tickets that were used to determined money payoffs. After the decisions for each matching determined everyone's point earnings, two 10-sided dice were thrown to determine a random number between 0 and 99, and the subject would earn a $2.00 payoff if the die throw was less than or equal to the point earnings, $.50 otherwise.[8] After being matched with all 6 respondents, the proponent and respondent roles were reversed for the final six matchings.[9]

The main feature of the experimental results for this game is that the human data converge to the intuitive outcome (I,C). The proportion of less refined, sequential outcomes (S,C) fell from .19 in the first 6 periods to .06 in the final 6 periods. Conversely, the proportion of the intuitive (I,C) outcomes increased from .47 to .67. The outcome proportions understate the strength of the intuitive equilibrium, since deviations were regularly punished. When an S signal was sent, it was met with the D response that supports the intuitive equilibrium in 18 out of 22 cases in the last six matchings. Banks, Camerer, and Porter (1994) also ran an experiment with game N (including a third signal that was never used). Although their procedures were somewhat different, their summary results essentially match ours, with the proportion of (I,C) outcomes increasing from .53 to .68 in the final half of their experiment.

---

[8] The points earned thus determine the probability of winning the high prize. This binary lottery procedure is intended to induce risk neutrality in points, since the expected utility of the money payoffs is linear in the points. For a discussion of this procedure and further references, see Davis and Holt (1993, pp. 472-6).

[9] The "no-contagion" rotation protocol prevented anyone from ever being matched with a previous partner, or with someone who had been matched someone who had been matched with them, etc. The objective of this deterministic matching procedure was to preserve the one-period nature of the game. The procedure cannot protect against the possibility that one player's actions influence those of previous partners in a later matching. Therefore, the results of two-person games taken from the same session are not independent observations. Brandts and Holt (1993) contains the instructions used and a more detailed description of procedures and matching protocols.

3. A SIMULATION MODEL OF BAYESIAN LEARNING AND ADJUSTMENT

Even though the (I,C) outcome predominated in the game N data, the type B proponents started off sending I only about one-fourth of the time, and the type A proponents chose I about three-fourths of the time. Those who sent the I signal tended to get the C (reward) response, and those who sent the S signal got a high fraction of D (punishment) responses, which eventually caused them to switch to I. This pattern in the adjustment process is shown in figure 1, where the data points are averages for adjacent pairs of matchings: 1-2, 3-4, etc. In our judgement, the combination of adjacent observations yields a clearer representation of behavior over time than period-by-period data. The observed pattern, which is not addressed by the refinements theory, caused us to consider whether the equilibrium selected is due to the path of adjustment rather than to the logic of the theoretical refinements.

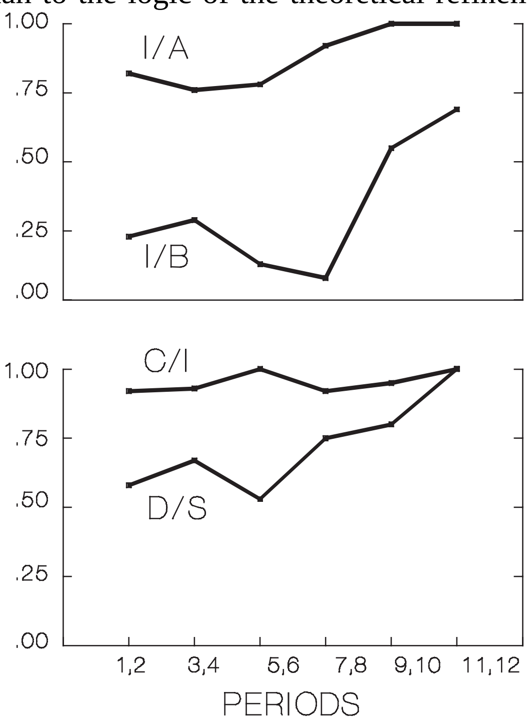

Figure 1. Game N, Human Data

The pattern of adjustment that leads to the preponderance of the intuitive (I,C) outcomes could be due to a naive reasoning process like the following. The top row of payoffs for game N in table 1 shows that the sum of possible payoffs for a type A proponent sending message I is higher than for message S, whereas the reverse is true for a type B proponent. If the proponent initially has a flat prior, believing that each response is equally likely regardless of signal, then the expected payoff for each signal is the sum of three possible payoffs divided by three. Therefore, initial choices by the proponent could lead to “type dependence”, with type A sending message I and type B sending message S. Similarly, a naive respondent would initially believe that a message is equally likely to have come from either type. These flat priors will lead to the C response to either message, since this response provides the maximum column sum. This C response should not alter the relation between proponents’ types and messages, since each type is getting its maximum payoff of 45.

But as respondents discover the type dependence (I from type A and S from type B), they will change their response to S from the reward, C, to the punishment, D. This response change in turn causes the type B proponents to switch to signal I, especially if they have previously been a type A who sent the I signal and got the C response.

The above description of how subjects gain experience with the game is admittedly imprecise, since the pattern of proponent types is random from matching to matching and not every subject gets experience with both types. The simulation program that is explained next is a more careful representation of our intuition about how subjects make decisions. The prior probabilities for each simulated subject were set to be uniform. For proponents, this means that the probability is 1/3 for each of the three responses to I, and 1/3 for each of the three responses to signal S. For respondents, the prior probabilities that either signal came from a type A proponent were initialized to 1/2. In the first matching, the six simulated proponents are matched with the six simulated respondents, and each proponent is given a randomly determined type. The flat prior and the relevant payoffs can then be used to calculate the optimal signal (I for type A and S for type B). The respondents see the signal sent by the proponent with whom they are matched and use the flat priors to calculate their optimal responses (C to each signal).

The key to the simulation program is the use of observed decisions to update the proponents' probabilities of the response to each signal, and the respondents' probabilities of the proponent type conditional on the observed signal. If the prior is Dirichlet and the observation is a random draw from a multinomial distribution with these priors, then the posterior probabilities are found by adding one to the numerator of probability for the event observed and adding one to the denominator to the probabilities for all events (DeGroot, 1970, p.174). For example, a proponent who observes a C response to an I signal in the first matching would increase Pr(C|I) from 1/3 to (1+**1**)/(3+**1**), and would decrease Pr(D|I) and Pr(E|I) from 1/3 to 1/(3+**1**), where the bold type is used to represent the effect of the observation. Since a single observation adds 1 to the denominators and to the relevant numerator of these probabilities, the initial prior of 1/3 is equivalent to having observed one of each response before the start of the learning process.

The updating for the simulated respondents is done in an analogous manner: they use the mechanical updating rule to calculate the posteriors on signal given type, Pr(I|A) and Pr(I|B).

Then Bayes' rule and the fact that the types are equally likely are used to calculate the probabilities of type given signal: Pr(A|I) and Pr(A|S). The latter probabilities are then used to calculate the respondents' expected payoffs conditional on the observed signal. The exogenous randomness corresponding to proponents' types generally causes the posterior probabilities to evolve differently for each simulated players, and the program keeps track of the relevant sums.

One issue in the simulation is how to model beliefs when the subjects' roles change after the sixth matching. A player who has just switched roles has never encountered someone in the other role. We assume a player's expectations after a role change are based on that player's own decisions in the previous role. This assumption was implemented in the simulation by keeping track of each player's own decisions, and using the same Bayesian process to update "fictional" beliefs based on those decisions. These are the beliefs that would have been formed by someone who was looking over the player's shoulder period by period and trying to predict the next decision. When the player reverses roles, these fictional beliefs are used to form the player's actual prior beliefs in the new role. For example, a respondent who had always responded to I with C would have a high fictional Pr(C|I), which would become the simulated player's actual prior in matching 7 when the player becomes a proponent.

The results of simulations of game N are shown by the solid lines in figure 2 (ignore the dashed lines for now). The data represent an average across 50 simulations, grouped in adjacent periods for comparability with figure 1. The simulation results match the qualitative pattern in the human data in figure 1, with the D|S punishments rising before the type B proponents switch from S to I in the later matchings. Note that with human data, the D|S response is quite common in early periods, in contrast to the simulated data. The initial propensity to choose D after seeing S is an indication that some respondents may have anticipated the type dependence, i.e. that the S signal was more likely to be sent by a type B proponent.

The switch from S to I signals for the simulated type B proponents is quite slow, as shown by the I/B line in figure 2. One way to speed up the convergence is to downgrade the importance of initial prior beliefs. The Dirichlet prior allows us to reduce the importance of the prior relative to the data by specifying the priors as a ratio of α parameters: $\Pr(C|I) = \alpha/(\alpha+\alpha+\alpha)$, $\Pr(D|I) = \alpha/(\alpha+\alpha+\alpha)$, etc. These prior probabilities are still flat, but they change more quickly in response to decisions when $\alpha < 1$, since the observed decision is still coded as

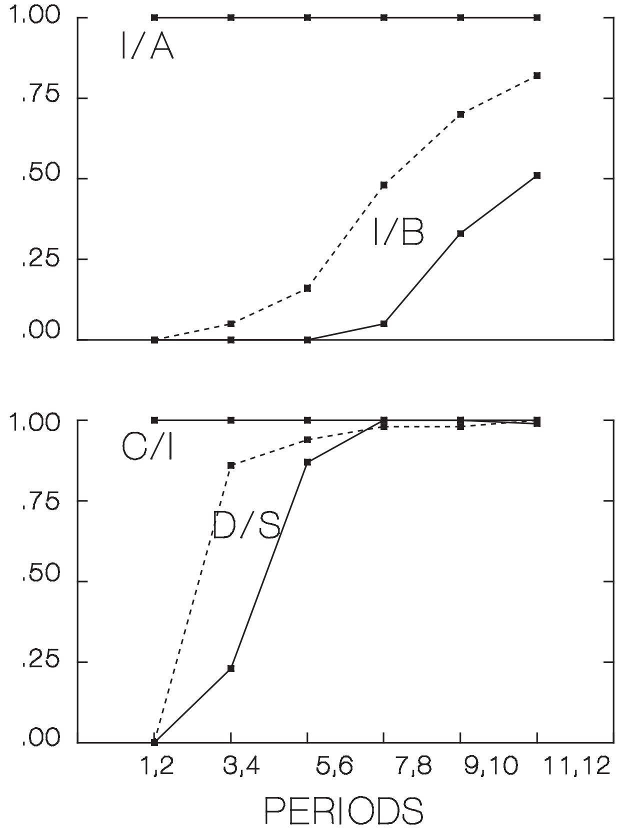

Figure 2. Game N, Simulated Data with Naive Bayesian Learning (solid lines: α = 1, dashed lines: α = .5)

1 and added to the numerator and denominator for the probability of the observed decision. For example, a C response to an I signal in the first matching would yield: $Pr(C|I) = (1+\alpha)/(1+\alpha+\alpha+\alpha)$, which exceeds 1/2 when $\alpha < 1$.

The dashed lines in figure 2 show results across fifty simulations where the value of α has been reduced from 1 to .5. The lines for I/A and C/I are the same for the two values of α and, therefore, overlap in the figure. The dashed lines for I/B and for D/S rise more quickly and tend to stay to the left of the solid lines. In relation to the human data shown in figure 1, the main thing to notice is that the B-types switch to the I signal earlier and more completely in the simulations with α = .5. Smaller values of α tend to increase the I/B and D/S proportions, especially in earlier periods, but at a strongly diminishing rate.

In summary, the simulations reported above track the main qualitative properties of the adjustment to the more refined equilibrium. There are, however, two features of the human data which are not well represented by the simulations shown above. First, the simulations yield adjustment paths which lack some of the variability of subjects' behavior. Second, values of the simulations corresponding to the D/S response are too low compared to the human data. Modifications of the simulation model that improve the correspondence with human data will be considered in section 5. First, we will show that the simple version of the simulation model can be used to predict human data that converge to the less refined equilibrium in a game to be considered next.

Table 2. Game R
(proponent's payoff, respondent's payoff)

	response				response		
	C	D	E		C	D	E
type A sends I	30, 30	0, 0	50, 35	type A sends S	45, 90	15, 15	100, 30
type B sends I	30, 30	30, 45	30, 0	type B sends S	45, 0	0, 30	0, 15

4. GAME R

Game R in table 2 has the same pure-strategy equilibrium configuration as game N, with a pooling equilibrium (signal S) that is sequential but not intuitive, and with another equilibrium (signal I) that is both sequential and intuitive.[10] The payoffs for this game have been chosen to induce a different adjustment process that reverses the initial correlation between signals and types. It follows from a naive belief in equally likely responses that message S is more attractive for a proponent of type A, and message I is more attractive for a type B. In this way, we reverse the correlation between messages and types that was predicted for naive players in game N. This reversal is the reason why we refer to this game as R in the discussion that follows, while game N refers to the game where type dependence is "normal".

Next, consider the initial decisions of naive respondents in game R. Given an initial belief that the signal contains no information about the proponent's type (and hence both types are considered equally likely), it is straightforward to show that the best response to each message is C. This naive response gives the type B proponent an incentive to switch from I to S, and it gives the type A proponent an incentive to stay with S. As a result, play would converge to the sequential equilibrium. Once play settles down on the (S,C) outcome, the

[10] There is also a mixed equilibrium for game R, with Pr(I|A) = .64, Pr(I|B) = .11, Pr(C|I) = .75, Pr(D|I) = 0, Pr(C|S) = .67, and Pr(D|S) = .33. The data to be discussed below do not show any resemblance to these predictions, and therefore, we will not discuss this equilibrium further. We are indebted to Richard McKelvey for pointing out the existence of this equilibrium.

respondent is likely to remember that an I message was previously sent by a type B player. Given this, the respondent will answer a deviation I message with D, which is precisely the best response if the deviant is expected to be of type B, a belief that is consistent with previous experience. Nevertheless, this belief is inconsistent with the intuitive criterion, since the type B player earns 45 in the sequential equilibrium, which is greater than any of the possible payoffs that could result from a deviation to the I message.

To summarize the experimental design, we structured games N and R so that a series of matchings with different players could generate different patterns of adjustment to equilibrium. In game N, the initial type dependence is for a naive type A to send message I and for a naive type B to send message S. This type dependence pattern should be reversed in game R.[11] Once behavior settles down, the past experience with out-of-equilibrium decisions determines the beliefs about what would happen off of the equilibrium path once an equilibrium is reached. Although these beliefs are consistent with the adjustment experience, they are inconsistent with the intuitive criterion in the case of game R.

[11] But in moving from game N to game R, the reversal of the adjustment process necessitated changes in the equilibrium payoffs. Unlike the case for game N, the sequential (S,C) outcome is better than the (I,C) outcome for both proponent types in game R. This observation led us to consider another game, with the same initial "reverse type dependence" as game R and the same equilibrium payoffs as game N. The data for this additional game, game 5 reported in Brandts and Holt (1993), also yielded data that were largely inconsistent with the intuitive criterion.

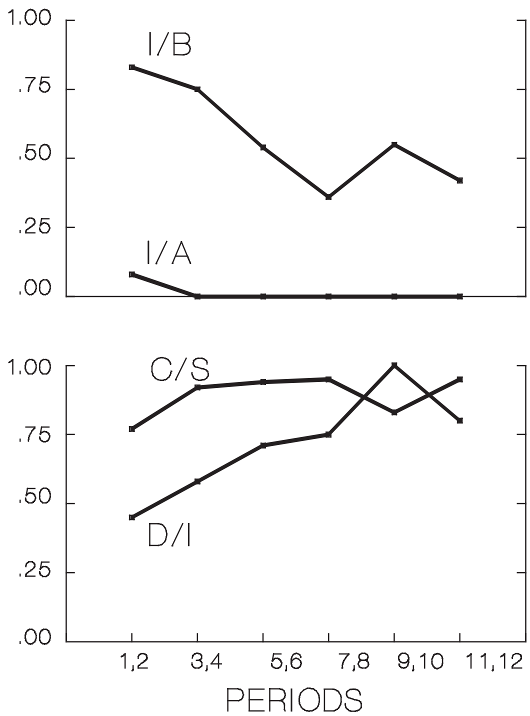

Figure 3. Game R, Human Data

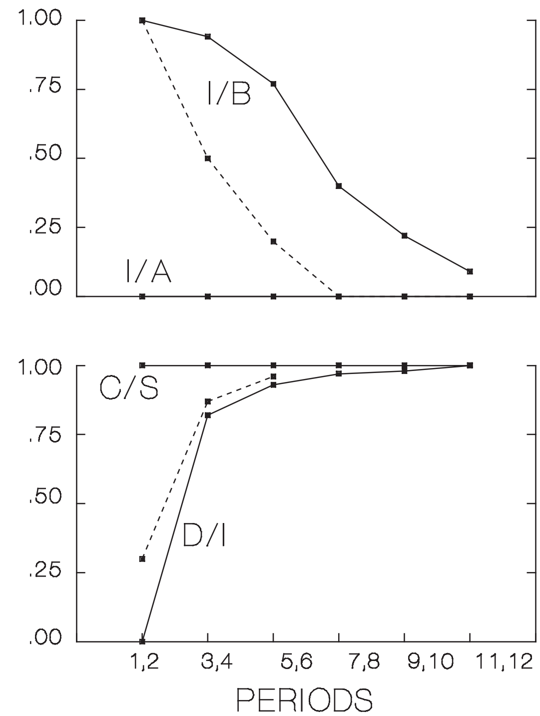

Figure 4. Game R, Simulated Data
Naive Bayesian Learning
(solid lines: $\alpha = 1$, dashed lines: $\alpha = .5$)

Figure 3 shows the adjustment path for the human data for game R, formatted as in figure 1. The initial reverse type dependence is quite clear. As predicted, the frequency of D/I increases in later periods and this leads to a preponderance of sequential (S,C) outcomes in the last half of the matchings.

The convergence of subjects' decisions to the less refined equilibrium in game R is about as strong as the convergence to the more refined equilibrium in game N. This can be seen from the summary human data reported on the left side of table 3. Outcomes are categorized as intuitive (I,C), sequential (S,C), and non-Nash, which are not shown. Recall that the proportion of intuitive outcomes in game N increases from about half (.47) in the first six periods to about two thirds (.67) in the final six periods. This pattern is reversed for game R, with the proportion of *sequential* outcomes increasing from .51 to .72. If the game N results are interpreted as supporting the intuitive criterion and stronger refinements, the results for game R are clearly unsatisfactory. In the final 2 periods, 79% of the messages are intuitive in game N, and 79% of

Table 3. Proportions of Outcomes by Refinement

	Human Data		Simulated Data*	
	Intuitive	Sequential	Intuitive	Sequential
Game N (periods 1-6)	.47	.19	.46	.29
Game N (periods 7-12)	.67	.06	.64	.07
Game R (periods 1-6)	.14	.51	.10	.78
Game R (periods 7-12)	.02	.72	.02	.93

* Data averaged over 50 simulations with parameters: $\alpha = .5$, $\lambda = .2$, and $s = .2$.

the messages are sequential in game R. Also, the 100% of the responses in the final 2 periods of game N are those specified by the intuitive equilibrium (C for I and D for S), and 92% of the responses in the final 2 periods of game R are those specified by the sequential equilibrium (C for S and D for I). Recall that a theoretical refinement is a way of *ruling out* a subset of the Nash equilibria. To be useful, such a refinement should only fail in rare and extreme cases.

Computer simulations for game R provide further support for the notion that it is the adjustment process that determines which equilibrium is determined. Figure 4 above shows the results across fifty simulations with game R. The simulated data converge to the less refined equilibrium, and the qualitative features of the adjustment path are similar to those for the human data in figure 3. The main point is that the naive Bayesian learning model predicts the underlying pattern of adjustment in this game. One obvious difference between the human and simulated data for game R is in the D/I lines.[12] The human respondents seem to anticipate the type dependence and punish the I signal that is likely to be sent by type B proponents in early periods. The incorporation of strategic elements into the simulation model is discussed in the next section.

[12] In figure 4, the dashed line for D/I stops after periods 5 and 6 due to the fact that no more I messages are sent in any of the simulations.

5. Decision Errors and Strategic Behavior

There are two features of the human data for games N and R which are not captured by the simulations of sections 3 and 4. First, the convergence in the simulated data for both games is cleaner than for the human. Observe, for instance, how the values of I/A and for C/I exhibit some variation for the human data in figures 1 and 3, while the corresponding values for the simulated data are constant at the equilibrium levels in all periods. Second, the use of the D "punishment" in periods 1 and 2 is more frequent for the human data than for the simulated data.

By adding decision errors and some strategic behavior, the simulation model will be shown to produce data that exhibit the two features just described. Decision errors may be the explanation for the variation in human data which is not present in the simulated data. Anticipation of proponent behavior by some respondents may explain why human data in both games exhibit a higher incidence of punishments than is observed in the simulated data.

The introduction of decision errors is done with a logistic distribution in which the probability of each decision depends on the expected payoffs of all of the alternative decisions.[13] For example, a respondent who sees signal I can use the posterior probabilities of the proponent type to calculate the expected payoffs for each of the three responses, which will be denoted: $E\pi_{C|I}$, $E\pi_{D|I}$, and $E\pi_{E|I}$. Then the probability for each decision is a ratio of exponential expressions, with the error parameter, λ:

$$\textit{Probability of choosing C given I} \;=\; \frac{e^{\lambda E\pi_{C|I}}}{e^{\lambda E\pi_{C|I}} + e^{\lambda E\pi_{D|I}} + e^{\lambda E\pi_{E|I}}}. \qquad (1)$$

As $\lambda \rightarrow 0$, decisions become completely random since the decision probabilities go to 1/3 for each response (1/2 for each signal), irrespective of the expected payoffs. On the other hand, errors vanish as $\lambda \rightarrow \infty$, and the decision with the highest expected payoff is selected with probability 1.

Figure 5 shows the average data for 50 simulations of game N with decision errors (λ = .3) and some underweighting of the prior (α = .5). A comparison with the human data in figure 1 shows that the fit is quite good, except for the fact that humans appear to be more strategic in

---

[13] The addition of decision errors will keep the simulations from ever settling down on a pure strategy equilibrium in which all probabilities are zero or one. As learning occurs and the probabilities stabilize in our simulations, the result will be a type of stochastic equilibrium, like that proposed by McKelvey and Palfrey (1995).

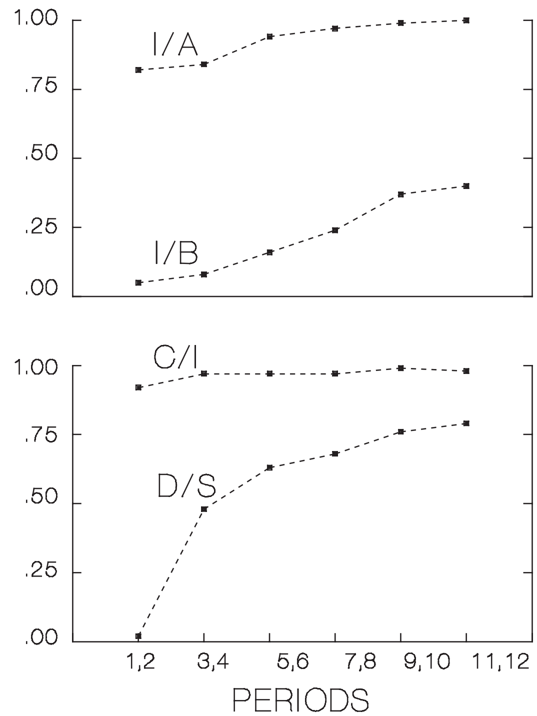

Figure 5. Game N, Simulated Data
Naive Bayesian Learning (α = .5)
Logistic Decision Errors (λ = .3)

starting off with a high proportion of D responses to the S signal, in anticipation of type dependence.

The naive learning that is built into the simulations thusfar does not require that subjects act strategically; they just maximize with respect to relative frequencies of previously observed decisions. To think strategically, a subject must try to see the situation from the other player's point of view, and then anticipate the other player's decision. This requires some knowledge of the other player's payoffs. In contrast, the decisions made in the simulations reported thusfar do not make any use of other players' payoffs. Partow and Schotter (1993) provide an interesting test of naive learning models by replicating four of the five signaling games contained in Brandts and Holt (1993), with two modifications of the procedures. First, subjects were only told their own payoffs, thereby precluding strategic play. Second, subjects only interacted for six matchings, since role reversal would have provided information about payoffs for the other role. In general terms, their results support the notion that naive play is consistent with the behavior in the experiments.[14]

The way we add strategic anticipation to the simulation model is motivated by the following representation of the decision problem facing a player who must predict the behavior

[14] For game N, their results show a preponderance of sequential outcomes over intuitive ones, in contrast to our results for periods 1-6. A possible explanation of this difference is that, although type dependence in our experiments is similar to theirs, the D/S response is more frequent in ours. The rationale for this could be some degree of strategic anticipation by subjects. However, results from our simulation model show that the naive model with enough inertia in the priors (α = 1) can in periods 1-6 yield proportions of sequential and intuitive outcomes like the ones found by Partow and Schotter. Moreover, in these simulations data from periods 7-12 exhibit the higher proportion of intuitive outcomes of our experiments with game N. We conclude that the Partow and Schotter results are not inconsistent with our naive simulations.

of a partner who has never been encountered previously. A player might begin by reasoning: "At least I know that my partner has encountered a series of players in the same situation as myself. If these players behaved like I have up to now, what pattern of decisions would my partner have seen, and what would my partner's best response to that pattern be, given what I know about his or her payoffs?" Given this specific way of anticipating the other's action, the player will then choose his best choice.

In the simulation model, the fictional beliefs attributed to a player's partners are updated period by period. In the initial period, each player assumes that the other players has flat priors (on each response given the signal, or on each type given the signal). With strategic anticipation, for example, a proponent will expect a respondent to choose C in response to either signal, since this is the respondent's best choice given flat priors. In each subsequent period, each player calculates what the beliefs would be for another player who had observed all of the player's own decisions and used Bayes' rule to update. Given the assumption about the other player's current beliefs, a player can calculate the other player's expected payoffs for each decision, and use the logistic parameter to infer probabilities for each of the other player's decisions. We call these "strategic probabilities", since they are based on the best response for someone who has been previously matched with opponents who have behaved like oneself. Such strategic probabilities are calculated on the basis of a knowledge of the other player's payoffs, whereas naive Bayesian beliefs are calculated from observed frequencies of other players' decisions.

We decided to allow for different degrees of strategic anticipation, represented by different values of a parameter, s. The value of s determines the weight assigned to the strategic probabilities in the simulations. In particular, each subject's beliefs are calculated as (1-s) times the probabilities calculated from decisions made by one's own partners in previous matchings, plus (s) times the strategic probabilities calculated from the best response to one's own decisions in previous matchings. All simulations reported above have no strategic anticipation, i.e. s = 0. At the other extreme, with s = 1, a player essentially assumes that the other is making a best response to a forecast based on the player's own previous decisions. Even with s = 1 this type of strategic anticipation is, of course, rather limited. It is possible to model behavior as being much more forward looking, in an "infinite regress." After looking at the adjustment patterns in the human data, we decided to limit the extent of strategic anticipation in the simulation model by representing players as looking one step ahead.[15]

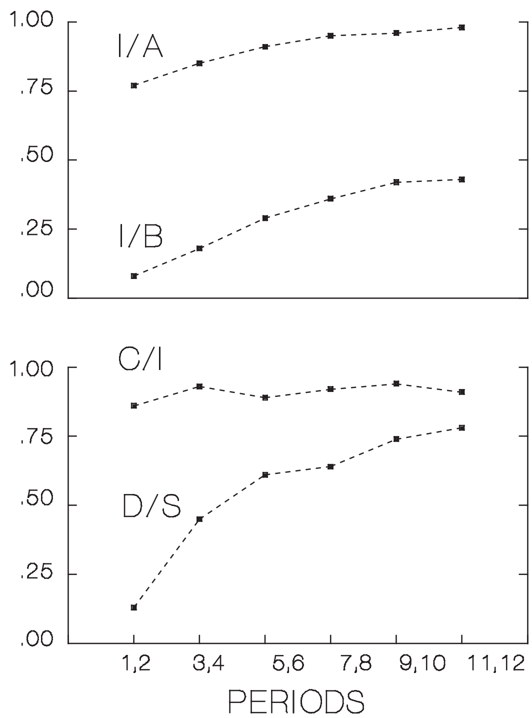

Figure 6. Game N, Simulated Data
Naive Bayesian Learning (α = .5)
Logistic Decision Errors (λ = .2)
Strategic Anticipation (s = .2)

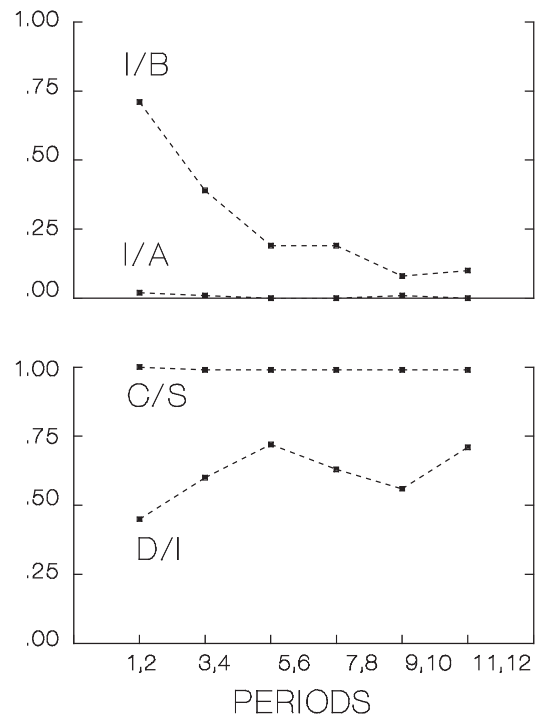

Figure 7. Game R, Simulated Data
Naive Bayesian Learning (α = .5)
Logistic Decision Errors (λ = .2)
Strategic Anticipation (s = .2)

We began by using the parameters from the simulations in figure 5, with (α = .5, λ = .3, s = 0), and raising the strategic parameter, s, to .2. The addition of strategic elements tends to make the convergence to equilibrium a little cleaner, so we increased the level of errors by reducing λ to .2. This simulation is shown in figure 6, and a simulation with the same parameters for game R is shown in figure 7. The addition of strategic anticipation raises the incidence of the D punishments in the first two periods, especially in game R. These simulations track the primary features of the adjustment pattern in the human data. The summary data are

[15] Stahl and Wilson (1994) report experimental evidence from 3x3 symmetric games that supports the notion of limited strategic anticipation.

categorized by outcome on the right side of table 3 above. The outcome proportions roughly correspond to those for the human data, with the notable exception that the simulation for game R exhibits a tighter convergence to the less refined (S,C) outcome.

6. CONCLUSIONS

The experiments and the simulations reported in this paper show how equilibrium selection in signaling games can be explained in terms of a specific adjustment model of naive Bayesian learning with decision errors and some strategic anticipation. Our analysis of human data is supported by computer simulations with the adjustment model. More specifically, our simulation results provide a good approximation of the human data, period by period in real time. The analysis identifies the behavioral basis for the fact that a number of refinements are not good predictors of outcomes in experimental games. Restrictions on beliefs that *a priori* seem reasonable starting at an equilibrium may not be appropriate when equilibria are approached through adjustment processes like the ones described above.

At a more general level, our results illustrate how stable economic situations can be seen as outcomes of adjustment processes. The absence of a widely accepted theory of adjustment to equilibrium is one of the rather unsettling aspects of economic theory. Theoretical models of out-of-equilibrium behavior have often been criticized as being *ad hoc* or arbitrary. Laboratory experiments like the ones presented in this paper provide data for the study of convergence to equilibrium. Theoretical insights can be tested with further experiments and with simulations. A simulation that fits the data for one experiment can be used to design another. If the fit is not good in certain respects, then the nature of the disparity can be used to modify the theoretical model on which the simulation is based. Simulation models that yielded satisfactory results could then be used to specify the functional form of an econometric learning model. This interactive process, which is common in the physical sciences, has a lot of promise in terms of bringing theory and observation closer together.

We believe that this interactive use of experiments, simulation models and theory can make a substantive contribution to economics. It can lead to an observationally based method for equilibrium selection in models with multiple equilibria. It can also lead to a prediction of behavior in environments that are not stationary long enough for full adjustment to any single equilibrium to take place. The relevance of intermediate range predictions has also been emphasized by Roth and Erev (1995).

Our main conclusion is that, as economists, we need to take the process of adjustment to equilibrium seriously. The results contained in this paper support the notion that adjustment theories should model how people actually learn and adapt when they are relatively unfamiliar with the environment. In this sense, we were influenced by the important work of Cyert and DeGroot (1987 and references therein), who devised models of sequential Bayesian learning for a wide array of economic applications. Like Cyert and DeGroot and many of their coauthors at Carnegie-Mellon, we find that computer simulations of Bayesian learning and adjustment are especially useful techniques when standard equilibrium assumptions do not offer adequate explanations of observed behavior patterns.

REFERENCES

Anderson, Lisa, 1994, Information cascades: A logistic error model of laboratory data, working paper, University of Virginia.

Ball, Sheryl and Roy Gardner, 1993, The evolution of behavior in experimental games, working paper, Indiana University.

Banks, Jeffrey S., Colin F. Camerer and David Porter, 1994, Experimental tests of Nash refinements in signaling games, *Games and Economic Behavior 6*, 1-31.

Brandts, Jordi and Charles A. Holt, 1992, An experimental test of equilibrium dominance in signaling games, *American Economic Review* 82, 1350-1365.

Brandts, Jordi and Charles A. Holt, 1993, Adjustment patterns and equilibrium selection in experimental signaling games, *International Journal of Game Theory* 22, 279-302.

Cooper, David J., Susan Garvin, and John H. Kagel, 1994, Adaptive learning versus equilibrium refinements in an entry limit pricing game, working paper, University of Pittsburgh.

Cho, In-Koo and David M. Kreps, 1987, Signaling games and stable equilibria, *Quarterly Journal of Economics* 102, 179-221.

Crawford, Vincent P., 1991, An 'evolutionary' interpretation of Van Huyck, Battalio, and Beil's experimental results on coordination, *Games and Economic Behavior* 3, 25-29.

Cyert, Richard M. and Morris H. DeGroot, 1987, *Bayesian analysis and uncertainty in economic theory* (Rowman & Littlefield, Totowa, NJ).

DeGroot, Morris H., 1970, *Optimal statistical decisions* (McGraw-Hill, New York).

Friedman, Daniel, 1991, Evolutionary games in economics, *Econometrica* 59, 637-666.

Friedman, Daniel, 1992, Evolutionary games: Some experimental results, working paper, University of California at Santa Cruz.

Gale, John, Kenneth G. Binmore and Larry Samuelson, 1995, Learning to be imperfect: The ultimatum game, *Games and Economic Behavior* 8, 56-90.

Harrison, Glenn W. and Kevin McCabe, 1992, Testing bargaining theory in experiments, in: Mark Isaac, ed., *Research in Experimental Economics*, vol. 5 (JAI Press, Greenwich, Conn.).

Holt, Debra J., 1993, Detecting learning in experimental games, working paper, Queens University.

Jordan, James S., 1991, Bayesian learning in normal form games, *Games and Economic Behavior* 3, 60-81.

Kalai, Ehud and Ehud Lehrer, 1993, Rational learning leads to Nash equilibrium, *Econometrica* 61, 1019-1045.

Kohlberg, Elon and Jean-Francois Mertens, 1986, On the strategic stability of equilibria, *Econometrica* 54, 1003-1037.

Mailath, George, 1992, Introduction: Symposium on evolutionary game theory, *Journal of Economic Theory* 57, 259-277.

Marimon, Ramon and Ellen McGrattan, 1995, On adaptive learning in strategic games, in: A. Kirman and M. Salmon, eds., *Learning and rationality in economics* (Basil Blackwell, Oxford) 61-101.

McCabe, Kevin, and Arijit Mukherji, 1993, An experimental study of learning in games, working paper, University of Minnesota.

McKelvey, Richard and Thomas R. Palfrey, 1995, Quantal response equilibria for normal form games, *Games and Economic Behavior* 10, 6-38.

Myers, S. and N. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187-221.

Partow, Zeinab and Andrew Schotter, 1993, Does game theory predict well for the wrong reasons: An experimental investigation, working paper, NYU.

Rassenti, Stephen J., Stanley S. Reynolds, Vernon L. Smith, and Ferenc Szidarovszky, 1993, Learning and adaptive behavior in repeated experimental Cournot games, working paper, University of Arizona.

Roth, Alvin E. and Ido Erev, 1995, Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term, *Games and Economic Behavior* 8, 164-212.

Selten, Reinhard and Rolf Stoecker, 1986, End behavior in sequences of finite prisoner's dilemma supergames, *Journal of Economic Behavior and Organization* 7, 47-70.

Sopher, Barry, and Dilip Mookherjee, 1993, An experimental study of learning and decision costs in constant-sum games, working paper, Rutgers University.

Spence, A. Michael, 1973, Job market signaling, *Quarterly Journal of Economics* 87, 335-374.

Stahl, Dale O. and Paul W. Wilson, 1994, On players' models of other players: Theory and experimental evidence, working paper 9406, Center for Applied Research in Economics, University of Texas at Austin.

Van Huyck, John B., and Raymond C. Battalio, 1993, Learning in coordination games, working paper, Texas A&M.

Biographic Sketches:

Jordi Brandts received his Ph.D. from the University of Pennsylvania in 1986 and is currently a Research Economist at the Institute of Economic Analysis (CSIC) in Barcelona. His research interests are in experimental economics and public economics. He has published articles in leading journals, including the American Economic Review, the Journal of Public Economics, and Games and Economic Behavior.

Charles A. Holt received his Ph.D. from Carnegie-Mellon University in 1977 and is presently Chairman and Merrill Bankard Professor of Economics at the University of Virginia. He has authored numerous articles and the monograph, *Experimental Economics* (with Doug Davis). His research interests are experimenal economics, industrial organization, and game theory. He has served as president of the Economic Science Association and vice president of the Southern Economics Association.