

## American Economic Association

---

Anomalous Behavior in a Traveler's Dilemma?

Author(s): C. Monica Capra, Jacob K. Goeree, Rosario Gomez, Charles A. Holt

Source: *The American Economic Review*, Vol. 89, No. 3 (Jun., 1999), pp. 678-690

Published by: [American Economic Association](#)

Stable URL: <http://www.jstor.org/stable/117040>

Accessed: 26/08/2010 08:59

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=aea>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*.

<http://www.jstor.org>

# Anomalous Behavior in a Traveler's Dilemma?

By C. MONICA CAPRA, JACOB K. GOEREE, ROSARIO GOMEZ, AND CHARLES A. HOLT\*

The notion of a Nash equilibrium has joined supply and demand as one of the two or three techniques that economists instinctively try to use first in the analysis of economic interactions. Moreover, the Nash equilibrium and closely related game-theoretic concepts are being widely applied in other social sciences and even in biology, where evolutionary stability often selects a subset of the Nash equilibria. Many people are uneasy about the stark predictions of the Nash equilibrium in some contexts where the extreme rationality assumptions seem implausible. Kaushik Basu's (1994) "traveler's dilemma" is a particularly convincing example of a case where the unrelenting logic of game theory is at odds with intuitive notions about human behavior. The story associated with the dilemma is that two travelers purchase identical antiques while on a tropical vacation. Their luggage is lost on the return trip, and the airline asks them to make independent claims for compensation. In anticipation of excessive claims, the airline representative announces:

We know that the bags have identical contents, and we will entertain any claim between \$2 and \$100, but you will each be reimbursed at an amount that equals the *minimum* of the two claims submitted. If the two claims differ, we will also pay a reward of \$2 to the person making the smaller claim and we will deduct a penalty of \$2 from the reimbursement to the person making the larger claim.

Notice that, irrespective of the actual value of the lost luggage, there is a unilateral incentive to "undercut" the other's claim. It follows from

this logic that the only Nash equilibrium is for both to make the minimum claim of \$2. As Basu (1994) notes, this is also the unique strict equilibrium, and the only rationalizable equilibrium when claims are discrete. When one of us recently described this dilemma to an audience of physicists, someone asked incredulously: "Is this what economists think the equilibrium is? If so, then we should shut down all economics departments."

The implausibility of the Nash equilibrium prediction is based on doubts that a small penalty and/or reward can drive claims all the way to an outcome that minimizes the sum of the players' payoffs. Indeed, the Nash equilibrium in a traveler's dilemma is independent of the size of the penalty or reward. Economic intuition suggests that behavior conforms closely to the Nash equilibrium when the penalty or reward is high, but that claims rise to the maximum level as the penalty/reward parameter approaches \$0.

This paper uses laboratory experiments to evaluate whether average claims are affected by (theoretically irrelevant) changes in the penalty/reward parameter. The laboratory procedures are described in Section I. The second and third sections contain analyses of aggregate and individual data. Section IV presents a learning model that is used to obtain maximum likelihood estimates of the learning and decision error parameters. The fifth section considers behavior in the final periods after most learning has occurred, i.e., when average claims stabilize and behavior converges to a type of noisy equilibrium which combines a standard logit probabilistic choice rule with a Nash-like consistency-of-actions-and-beliefs condition. The learning/adjustment and equilibrium models are complementary, and together they are capable of explaining some key features of the data. Section VI concludes.

## I. Procedures

The data were collected from groups of 9–12 subjects, with each group participating in a

\* Capra: Department of Economics, Washington and Lee University, Lexington, VA 24450; Goeree and Holt: Department of Economics, University of Virginia, 114 Rouss Hall, Charlottesville, VA 22901; Gomez: Department of Economics, University of Malaga, 29013 Malaga, Spain. This project was funded in part by the National Science Foundation (Grant Nos. SBR-9617784 and SBR-9818683). We wish to thank Peter Coughlan, Rob Gilles, Susana Cabrera-Yeto, Irene Comeig, Nadège Marchand, and two anonymous referees for suggestions.

series of traveler's dilemma games during a session that lasted about an hour and a half. This type of experiment had not been done before, and we felt that more would be learned from letting the penalty/reward parameter,  $R$ , vary over a wide range of values. Therefore, we used two high values of  $R$  (\$0.50 and \$0.80), two intermediate values of  $R$  (\$0.20 and \$0.25), and two low values of  $R$  (\$0.05 and \$0.10). The penalty/reward parameter alternated between high and low values in parts A and B of the experiment. For example, session 1 began with  $R = \$0.80$  in part A, which lasted for 10 periods. Then  $R$  was lowered to \$0.10 in part B.

Subjects were recruited from economics classes at the University of Virginia, with the promise that they would be paid a \$6 participation fee plus all additional money earned during the experiment. Individual earnings ranged from about \$24.00 to \$44.00 for a session. We began by reading the instructions for part A (these instructions are available from the authors on request). Although decisions were referred to as "claims," the earnings calculations were explained without reference to the context, i.e., without mentioning luggage, etc. In each period, subjects would record their claim on their decision sheets, which were collected and randomly matched (with draws of numbered ping-pong balls) to determine the "other's claim" and "your earnings," and the sheets were then returned. Claims were required to be any number of cents between and including 80 and 200, with decimals being used to indicate fractions of cents. Subjects only saw the claim decision made by the person with whom they were matched in a given period. They were told that part A would be followed by "another decision-making experiment" but were not given additional information about part B. The penalty/reward parameter was changed in part B, and random pairwise matchings were made as before. Part B lasted for 10 periods, except in the first two sessions where it lasted for 5 periods.

## II. Data

The part A data are summarized in Figure 1. Each line connects the period-by-period averages of the 9–12 subjects in each group. There is a different penalty/reward parameter for each cohort, as indicated by the labels on the

right. The data plots are bounded by horizontal dashed lines that show the maximum and minimum claims of 200 and 80. The Nash equilibrium prediction is 80 for all treatments. The two highest lines in Figure 1 plot the average claims for low reward/penalty parameters of 5 and 10 (cents). The first-period averages are close to 180, and they stay high in all subsequent periods, well away from the Nash equilibrium. The two lowest lines represent the average claims for the higher penalty/reward parameters of 50 and 80. Note that with these parameters, the average claims quickly fall toward the Nash equilibrium. For intermediate reward/penalty parameters of 20 and 25, the average claims level off at about 120 and 145 respectively. The averages in the last five periods are clearly inversely related to the magnitude of the penalty/reward parameter, and the null hypothesis of no relation can be rejected at the 1-percent level.<sup>1</sup>

For some sessions, the switch in treatments between parts A and B caused a dramatic change in behavior. In the two sessions using  $R = 80$  and  $R = 10$ , for example, the behavior is approximately reversed, as shown in Figure 2.<sup>2</sup> There is some evidence of a sequence effect, since the average claims were higher for  $R = 10$  when this treatment came first than when it followed the  $R = 80$  treatment that "locked" onto a Nash equilibrium. In fact, the sequence effect was so strong in one session, with a treatment switch from  $R = 50$  to  $R = 20$ , that the data did not rise in part B after converging to the Nash outcome in part A. In all other sessions, the high- $R$  treatment resulted in lower average claims, as shown in Table 1.

Consider again the null hypothesis of no treatment effect, under which higher average claims are equally likely in both treatments. The alternative hypothesis is that average claims are higher

<sup>1</sup> Of the 720 (=6!) ways that the 6 session averages could have been ranked, there are only 6 possible outcomes that are as extreme as the one observed (i.e., with zero or one reversals between adjacent  $R$  values). Under the null hypothesis the probability of obtaining a ranking this extreme is: 6/720, so the null hypothesis can be rejected (one-tailed test) at the 1-percent level.

<sup>2</sup> Obviously, the part B data have not settled down yet after 5 periods, and therefore we decided to extend part B to 10 periods in subsequent sessions.

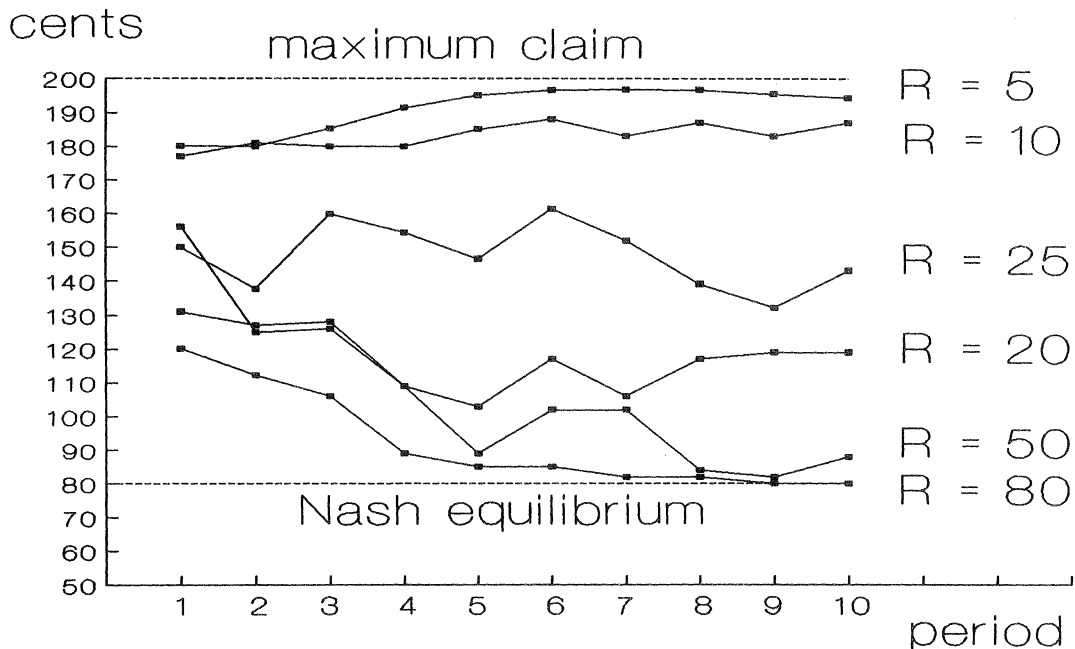


FIGURE 1. DATA FOR PART A FOR VARIOUS VALUES OF THE REWARD/PENALTY PARAMETER

for the treatments with a low value of  $R$ . The null hypothesis can be rejected at a 3-percent significance level using a standard Wilcoxon (signed-rank) nonparametric test. Thus the treatment effect is significant, even though it does not affect the Nash equilibrium.

Basically, the Nash equilibrium provides good predictions for high incentives ( $R = 80$  and  $R = 50$ ), but behavior is quite different from the Nash prediction under the treatments with low and intermediate values of  $R$ . In particular, as shown in Figure 1, the data for the low- $R$  treatments is concentrated at the *opposite* end of the range of feasible decisions. Basu's (1994) presentation of the traveler's dilemma involved low incentives relative to the range of choices so, in this sense, the intuition behind the dilemma is confirmed.<sup>3</sup> To summarize, the Nash

*equilibrium prediction of 80 for all treatments fails to account for the most salient feature of the data, the intuitive inverse relationship between average claims and the parameter that determines the relative cost of having the higher claim.*

Since the Nash equilibrium works well in some contexts, what is needed is not a radically different alternative, but rather, a generalization that conforms to Nash predictions in some situations (e.g., with high- $R$  values) and not in others. In addition, it would be interesting to consider dynamic theories to explain the patterns of adjustment in initial periods when the data have not yet stabilized. Many adjustment theories in the literature are based on the idea of movement toward a best response to previously observed decisions. The next section evaluates some of these adjustment theories and shows that they explain a high proportion of the directions of changes

<sup>3</sup> Basu (1994) does not claim to offer a resolution of the paradox, but he suggests several directions of attack. Loosely speaking, these approaches involve restricting individual decisions to sets,  $T_1$  and  $T_2$  for players 1 and 2 respectively, where each set contains all best responses to claims in the other person's set. Such sets may exist if open sets are allowed, or if some notion of "ill-defined category"

is introduced. Without further refinement these approaches do not predict the effects of the penalty/reward parameter on claim levels.

cents

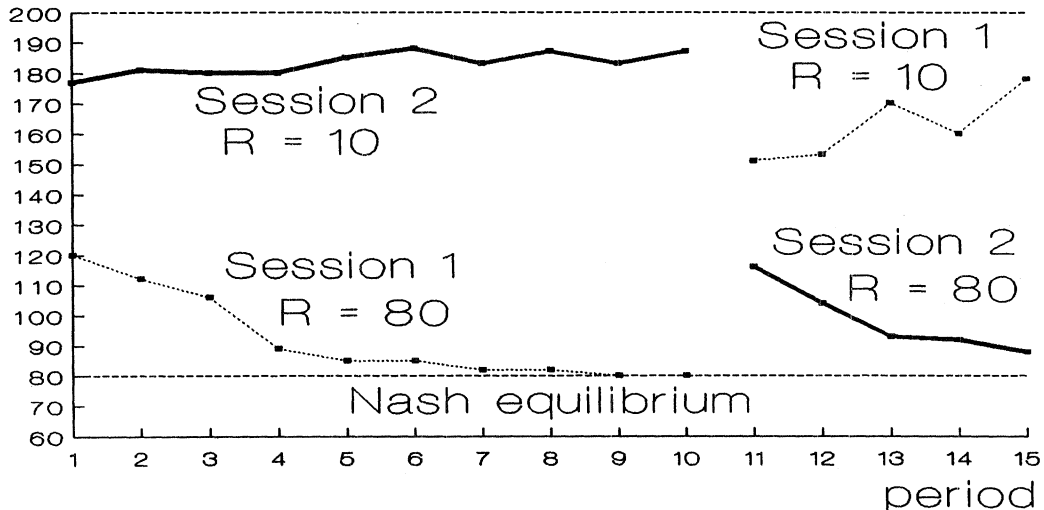


FIGURE 2. AVERAGE CLAIMS FOR PARTS A AND B OF SESSION 1 (DARK LINE) AND SESSION 2 (DASHED LINE)

in individual claims, but not the strong effect of the penalty/reward parameter on the levels of average claims. Then in Sections IV and V we present both dynamic and equilibrium theories that are sensitive to the magnitude of the penalty/reward parameter.

### III. Patterns of Individual Adjustment

One approach to data analysis is based on the perspective that people react to previous experience via what is called reinforcement learning in the psychology literature. In this spirit, Reinhard Selten and Joachim Buchta (1994) consider a model of directional adjustment in response to immediate past experience. The prediction is that changes are more likely to be made in the direction of what would have been a best response to others' decisions in the previous period. The predictions of this "learning direction theory" are, therefore, qualitative and probabilistic. The theory is useful in that it provides the natural hypothesis that changes in the "wrong" direction are just as likely as changes in the "right" direction. This null hypothesis is decisively rejected for data from auctions (Selten and Buchta, 1994).

To evaluate learning direction theory, we cat-

egorize all individual claims after the first period as either being consistent with the theory, "+", or inconsistent, "-". Excluded from consideration are cases of no change, irrespective of whether or not these are Nash equilibrium decisions. These cases of no change are classified as "na" (for not applicable). Table 2 shows the data classification counts by treatment. The (79, 21, 80) entry under the  $R = 5$  column heading, for example, means that there were 79 "+" classifications, 21 "-" classifications, and 80 "na" classifications. The percentage given just below this entry indicates that 79 percent of the "+" and "-" changes were actually "+". The percentages exclude the "na" cases from the denominator. The "percentage of +" row indicates that significantly more than half of the classifications were consistent with the learning direction theory. Therefore, the null hypothesis of no difference can be rejected in all treatments, as indicated by the " $p$ -value" row.

Note, however, that at least part of the success of learning direction theory may be due to a statistical artifact if subjects' decisions are draws from a random distribution, as described for instance by the equilibrium model in Section V. With random draws, the person who has the lower claim in a given period is more likely to be near the bottom of

TABLE 1—AVERAGE CLAIMS IN THE LAST FIVE PERIODS FOR ALL SESSIONS

Session	1	2	3	4	5	6
High- $R$ treatment	82	99	92	82	146	170
Low- $R$ treatment	162	186	86	116	171	196

the distribution, and hence to draw a higher claim in the next period. Similarly, the person with the higher claim is more likely to draw a lower claim in the following period. In fact, it can be shown that if claims are drawn from *any* stationary distribution, the probability is  $\frac{2}{3}$  that changes are in the direction predicted by learning direction theory.<sup>4</sup> But even the null hypothesis that the fraction of predicted changes is  $\frac{2}{3}$  is rejected by our data at low levels of significance.

One feature of learning direction theory in this context is that noncritical changes in the penalty/reward parameter  $R$  do not change the directional predictions of the theory. This is because  $R$  affects the magnitude of the incentive to change one's claim, but not the direction of the best response. This feature is shared by several other directional best-response models of evolutionary adjustment that have been proposed recently, admittedly in different contexts. For example, Vincent P. Crawford (1995) considered an evolutionary adjustment mechanism for coordination games that was operationalized by assuming that individuals switch to a weighted average of their previous decision and the best response to all players' decisions in the previous period. This adaptive learning model, which explains some key elements of adjustments in coordination game experiments, is similar to directional learning with the extent of directional movements determined by the relative weights placed on the previous decision and on the best response in the adjustment function. Another evolutionary for-

mulation that is independent of the magnitude of  $R$  is that of imitation models in which individuals are assumed to copy the decision of the person who made the highest payoff. With two-person matchings in a traveler's dilemma, the high-payoff person is always the person with the lower claim, regardless of the  $R$  parameter, so that imitation (with a little exogenous randomness) will result in decisions that are driven to near-Nash levels.<sup>5</sup> *To conclude, individual changes tend to be in the direction of a best response to the other's action in the previous period, but the strong effect of the penalty/reward parameter on the average claims cannot be explained by directional learning, adaptive learning (partial adjustment to a best response), and imitation-based learning models.*

#### IV. A Dynamic Learning Model with Logit Decision Error

In this section we present a dynamic model in which players use a simple counting rule to update their (initially diffuse) beliefs about others' claims. The modeling of beliefs is important because people will wish to make high claims if they come to expect that others will do the same. Although the only set of internally consistent beliefs and *perfectly rational* actions is at the unique Nash equilibrium claim of 80 for all values of  $R$ , the costs of increasing claims above 80 depend on the size of the penalty/reward parameter. For small values of  $R$  such deviations are relatively costless and some noise in decision-making may result in claims that are well above 80. As subjects encounter higher claims, the (noisy) best responses to expected claims may become even higher. In this manner, a relatively small amount of noise may

<sup>4</sup> Suppose that a player's draw,  $x_1$ , is less than the other player's draw,  $y$ . Then the probability that a next draw,  $x_2$ , is higher than  $x_1$  is given by:  $P[x_2 > x_1 | y > x_1] = P[x_2 > x_1, y > x_1] / P[y > x_1]$ . The numerator is equal to the probability that  $x_1$  is the lowest of three draws, which is  $\frac{1}{3}$ , and the denominator is equal to the probability that  $x_1$  is the lowest of 2 draws, which is  $\frac{1}{2}$ . So the relevant probability is  $\frac{2}{3}$ .

<sup>5</sup> Paul Rhode and Mark Stegeman (1995) and Fernando Vega-Redondo (1997) have shown that this type of imitation dynamic will drive outputs in a Cournot model up to the Walrasian levels.

TABLE 2—CONSISTENCY OF CLAIM CHANGES WITH LEARNING DIRECTION THEORY

	<i>R</i> = 5	<i>R</i> = 10	<i>R</i> = 20	<i>R</i> = 25	<i>R</i> = 50	<i>R</i> = 80	All treatments
Numbers of +, -, na	79, 21, 80	62, 15, 43	50, 7, 123	94, 23, 63	65, 12, 103	50, 3, 87	400, 81, 499
Percentage of + <sup>a</sup>	79	81	88	80	84	94	83
<i>p</i> -value <sup>b</sup>	<0.00003	<0.00003	<0.00003	<0.00003	<0.00003	<0.00003	<0.00003

<sup>a</sup> The percentage of "+" calculations excluded the nonapplicable "na" cases.

<sup>b</sup> Denotes the *p*-value for a one-tailed test.

move claims well above the Nash prediction when beliefs evolve endogenously over time in a sequence of random matchings.

This section begins with a relatively standard experience-based learning model.<sup>6</sup> There is clearly noise in the data, even in the final periods for some treatments, so for estimation purposes it is necessary to introduce a stochastic element. Noisy behavior is modeled with a probabilistic choice rule that specifies the probabilities of various decisions as increasing functions of the expected payoffs associated with those decisions. For low values of *R*, the "mistakes" in the direction of higher claims will be relatively less costly and, hence, more probable. Our analysis will be based on the standard logit model for which decision probabilities are proportional to exponential functions of expected payoffs. The logit model is equivalent to assuming that expected payoffs are subjected to shocks that have an extreme value distribution. These errors can be interpreted either as unobserved random changes in preferences or as errors in responding to expected payoffs.<sup>7</sup> The

logit formulation is convenient in that it is characterized by a single error parameter, which allows the consideration of perfect rationality in the limit as the error parameter goes to zero. Maximum likelihood techniques will be used to estimate the error and learning parameters for the data from the traveler's dilemma experiment.

To obtain a tractable econometric learning model, the feasible range of claims (between 80 and 200) is divided into  $n = 121$  intervals or categories of a cent. A choice that falls in category *j* corresponds to a claim of  $80 + j - 1$  cents. Players' initial beliefs, prior to the first period, are represented by a uniform distribution with weights of  $1/n$ . Hence, all categories are thought to be equally likely in the first period.<sup>8</sup> Players update their beliefs using a simple counting rule that is best explained by assigning weights to all categories. Let  $w_i(j, t)$  denote the weight that player *i* assigns to category *j* in period *t*. If, in period *t*, player *i* observes a rival's price that falls in the *m*th category, player *i*'s weights are updated as follows:  $w_i(m, t + 1) = w_i(m, t) + \rho$ , while all

<sup>6</sup> This model builds on the work of Jordi Brandts and Holt (1996), David J. Cooper et al. (1997), Dilip Mookherjee and Barry Sopher (1997), Yan Chen and Fang-Fang Tang (1998), Drew Fudenberg and David K. Levine (1998), Colin Camerer and Teck-Hua Ho (1999), and others.

<sup>7</sup> R. Duncan Luce (1959) provides an alternative, axiomatic derivation of this type of decision rule; he showed that if the ratio of probabilities associated with any two decisions is independent of the payoffs for all other decisions, then the choice probability for decision *i* can be expressed as a ratio:  $u_i / \sum_j u_j$ , where  $u_i$  is a "scale value" number associated with decision *i*. When scale values are functions of expected payoffs, and one adds the assumption that choice probabilities are unaffected by adding a constant to all expected payoffs, then it can be shown that the scale values are *exponential* functions of expected payoffs. There-

fore, any ratio-based probabilistic choice rule that generalizes the exponential form would allow the possibility that decision probabilities would be changed by adding a constant to all payoffs. While there is some experimental evidence that multiplicative increases in payoffs reduce noise in behavior (Vernon L. Smith and James M. Walker, 1997), we know of no evidence that behavior is affected by additive changes, except when subtracting a constant converts some gains into (focal) losses.

<sup>8</sup> Alternatively, the first-period data can be used to estimate initial beliefs. The assumption of a uniform prior is admittedly a simplification, but allows us to explain why the penalty/reward parameter has a strong effect even on decisions in the first period.

other weights remain unchanged. These weights translate into belief probabilities,  $P_i(j, t)$ , by dividing the weight of each category by the sum of all weights. The model is one of "fictitious play" in which a new observation is weighted by a learning parameter,  $\rho$ , that determines the importance of a new observation relative to the initial prior. A low value of  $\rho$  indicates "conservative" behavior in that new information has little impact on a player's beliefs, which are mainly determined by the initial prior.

Once beliefs are formed, they can be used to determine the expected payoffs of all the options available. Since in our model each player chooses among  $n$  possible categories, the expected payoffs are given by the sum

$$(1) \quad \pi_i^e(j, t) = \sum_{k=1}^n \pi(j, k) P_i(k, t), \quad j = 1, \dots, n,$$

where  $\pi_i(j, k)$  is player  $i$ 's payoff from choosing a claim equal to  $j$  when the other player claims  $k$ .

In a standard model of best-reply dynamics, a player simply chooses the category that maximizes the expected payoff in (1). However, as we discussed above, the adjustments in such a model will be independent of the magnitude of the key incentive parameter,  $R$ . We will therefore allow players to make nonoptimal decisions, or "mistakes," with the probability of a mistake being inversely related to its severity. The specific parameterization that we use is the logit rule, for which player  $i$ 's decision probabilities,  $D_i(j, t)$ , are proportional to an exponential function of expected payoffs:

$$(2) \quad D_i(j, t) = \frac{\exp(\pi_i^e(j, t)/\mu)}{\sum_{k=1}^n \exp(\pi_i^e(k, t)/\mu)}.$$

The denominator ensures that the choice probabilities add up to 1, and  $\mu$  is an error parameter that determines the effect of payoff differences on choice probabilities. When  $\mu$  is small, the decision with the highest payoff is very likely to be selected, whereas all decisions become equally likely (i.e., behavior becomes purely random) in the limit as  $\mu$  tends to infinity. To

summarize the key ingredients of our dynamic model: (i) players start with a uniform prior and use a simple counting rule to update their beliefs; (ii) these beliefs determine expected payoffs by (1); and (iii) the expected payoffs in turn determine players' choice probabilities by (2).<sup>9</sup>

This "logit learning model" can be used to estimate the error parameter,  $\mu$ , and the learning parameter,  $\rho$ . Recall that the probability that player  $i$  chooses a claim in the  $j$ th category in period  $t$  is given by  $D_i(j, t)$ , and the likelihood function is simply the product of the decision probabilities of the actual decisions made for all subjects and all 10 periods. The maximum likelihood estimates of the error and learning parameters of the dynamic learning model are:  $\mu = 10.9$  (0.6) and  $\rho = 0.75$  (0.12), with standard errors shown in parentheses. The error parameter is significantly different from the value of zero implied by perfect rationality, which is not surprising in light of the clear deviations from the Nash predictions.<sup>10</sup> If the learning parameter were equal to 1.0, each observation of another person's decision would be as informative as prior information, so a value of 0.7 means that the prior information is

<sup>9</sup> An alternative approach would specify that the probability of a given decision is an increasing function of payoffs that have been earned when that decision was made in the past. Thus high-payoff outcomes are "reinforced." See Alvin E. Roth and Ido Erev (1995) and Erev and Roth (1998) for a simulation-based analysis of reinforcement models in other contexts.

<sup>10</sup> We also estimated the learning model for each session separately, and in all cases the error parameter estimates were significantly different from zero, except for the  $R = 20$  session where the program did not converge. Recall that this treatment was the only one with an average claim that was out of the order that corresponds to the magnitude of the  $R$  parameter. The error parameter estimates (with standard errors) for  $R = 5, 10, 25, 50$ , and  $80$  were  $6.3$  (1.0),  $4.0$  (1.0),  $16.7$  (5.1),  $6.8$  (0.9), and  $9.5$  (0.7) respectively. These estimates are of approximately the same magnitude, but some of the differences are statistically significant at normal levels, which indicates that the learning model does not account for all of the "cohort effects." These estimates are, however, of roughly the same magnitude as those we have obtained in other contexts. Capra et al. (1998) estimate an error parameter of  $8.1$  in an experimental study of imperfect price competition. The estimates for the Lisa Anderson and Holt (1997) information cascade experiments imply an error parameter of about  $12.5$  (when payoffs are measured in cents as in this paper). Richard D. McKelvey and Thomas R. Palfrey (1998) use the Brandts and Holt (1996) signaling game data to estimate  $\mu = 10$  (they report  $1/\mu = 0.1$ ).



TABLE 3—PREDICTED AND ACTUAL AVERAGE CLAIMS

	$R = 5$	$R = 10$	$R = 20$	$R = 25$	$R = 50$	$R = 80$
Average claim in period 1	180	177	131	150	155	120
Average claim for periods 8–10	195	186	119	138	85	81
Average simulated claim for period 1	171	166	156	150	118	97
Average simulated claims for periods 8–10	178	170	155	148	107	85
Logit equilibrium prediction	183	174	149	133	95	88
Nash equilibrium prediction	80	80	80	80	80	80

slightly stronger than the information conveyed in a single observation.

Table 3 shows the relationship between average claims and simulation-based predictions of the logit learning model. The first row shows average claims observed in the first period of the experiment, where claims are highest for  $R = 5$  and  $R = 10$ , and lowest for  $R = 80$ . The second row shows the average observed claims for the final three periods; we see that claims rise slightly for the two low- $R$  treatments and fall for the high- $R$  treatments. The third row shows the first-period predictions of the dynamic model, based on the estimated error rate and the assumption of uniform initial priors. These predictions are also inversely related to the level of  $R$ . The predictions of the dynamic model for periods 8–10, shown in the fourth row, are obtained by letting a computer program keep track of 10 cohorts of 10 simulated subjects which begin with flat priors, make error-prone decisions, “see” the other’s decision, and update beliefs before being rematched randomly with another simulated subject. The simulated claims also show a tendency for claims to increase for low- $R$  values and decrease for high- $R$  values, but the treatment effect is a little too flat relative to the actual data. *To summarize, the parameter estimates for the logit learning model can be used in simulations to reproduce the qualitative features of observed adjustment patterns and the inverse relationship between the penalty/reward parameter and average claims.*

### V. A Logit Equilibrium Analysis

As players gain experience during the experiment, the prior information becomes consider-

ably less important. With more experience, there are fewer surprises on average, and this raises the issue of what happens if decisions stabilize, as indicated by the relatively flat trends in the final periods of part A for each treatment in Figure 1. An equilibrium is a state in which the beliefs reach a point where the decision distributions match the belief distributions, which is the topic of this section. Recall that in the previous section’s logit learning model, player  $i$ ’s belief probabilities,  $P_i(j, t)$  for the  $j$ th category in period  $t$ , are used in the probabilistic choice function (2) to calculate the corresponding choice probabilities,  $D_i(j, t)$ . A symmetric *logit equilibrium* is a situation where all players’ beliefs have stabilized at the same distributions, so that we can drop the  $i$  and  $t$  arguments and simply equate the corresponding decision and belief probabilities:  $D_i(j, t) = P_i(j, t) = P(j)$  for decision category  $j$ . In such an equilibrium, the equations in (2) determine the equilibrium probabilities (McKelvey and Palfrey, 1995, 1998).<sup>11</sup> The probabilities that solve these equations will, of course, depend on the penalty/reward parameter, which is desirable given the fact that this parameter has such a strong effect on the levels at which claims stabilize in the experiments. The equilibrium probabilities will also depend on the error parameter in (2), which can be estimated

<sup>11</sup> The logit equilibrium has been used to explain deviations from Nash behavior in some matrix games (Robert W. Rosenthal, 1989; McKelvey and Palfrey, 1995; Jack Ochs, 1995), in other games with a continuum of decisions, e.g., the “all-pay” auction (Simon P. Anderson et al., 1998a), public-goods games (Anderson et al., 1998b), and price-choice games (Gladys Lopez, 1995; Capra et al., 1998).

as before by maximizing the likelihood function. Instead of being determined by learning experience, beliefs are now determined by equilibrium consistency conditions.<sup>12</sup>

It is clear from the data patterns in Figure 1 that the process is *not* in equilibrium in the early periods, as average claims fall for some treatments and rise for others. Therefore, it would be inappropriate to estimate the logit equilibrium model with all data as was done for the logit learning model. We used the last three periods of data to estimate  $\mu = 8.3$  with a standard error of 0.5. This error parameter estimate for the equilibrium model is somewhat lower than the estimate for the logit learning model (10.9). This difference may be due to the fact that the learning model was estimated with data from all periods, including the initial periods where decisions show greater variability. Despite the difference in the treatment of beliefs, the logit learning and equilibrium models have similar structures, and are complementary in the sense that the equilibrium corresponds to the case where learning would stop having much effect, i.e., where decision and belief distributions are identical.

Once the error and penalty/reward parameters are specified, the logit equilibrium equations in (2) can be solved using *Mathematica*. Figure 3 shows the equilibrium probability distributions for all treatments with  $\mu = 8.3$ .<sup>13</sup>

<sup>12</sup> A theoretical analysis of the effect of  $R$  on equilibrium claims can be based on a continuous formulation in which probabilities are replaced by a continuous density,  $f(x)$ , with a distribution function,  $F(x)$ . In equilibrium, these represent players' beliefs about others' claims, which determine the expected payoff from choosing a claim of  $x$ , denoted by  $\pi^e(x)$ . The expected payoffs determine the claim density via a continuous logit choice function:  $f(x) = k \exp(\pi^e(x)/\mu)$ , where  $k$  is a constant of integration. This is not a closed-form solution for  $f(x)$ , since the claim distribution affects the expected payoff function. Nevertheless, it is possible to derive a number of theoretical properties of the equilibrium claim distribution. Anderson et al. (1998c) consider a class of auction-like games that includes the traveler's dilemma as a special case. For this class, the logit equilibrium exists, and is unique and symmetric across players. Moreover, it is shown that, for any  $\mu > 0$ , an increase in the  $R$  parameter results in a stochastic reduction in claims, in the sense of first-degree stochastic dominance.

<sup>13</sup> The density for  $R = 80$  seems to lie below that for  $R = 50$ . What the figure does not show is that the density for  $R = 80$  puts most of its mass at claims that are very close to 80 and has a much higher vertical intercept.

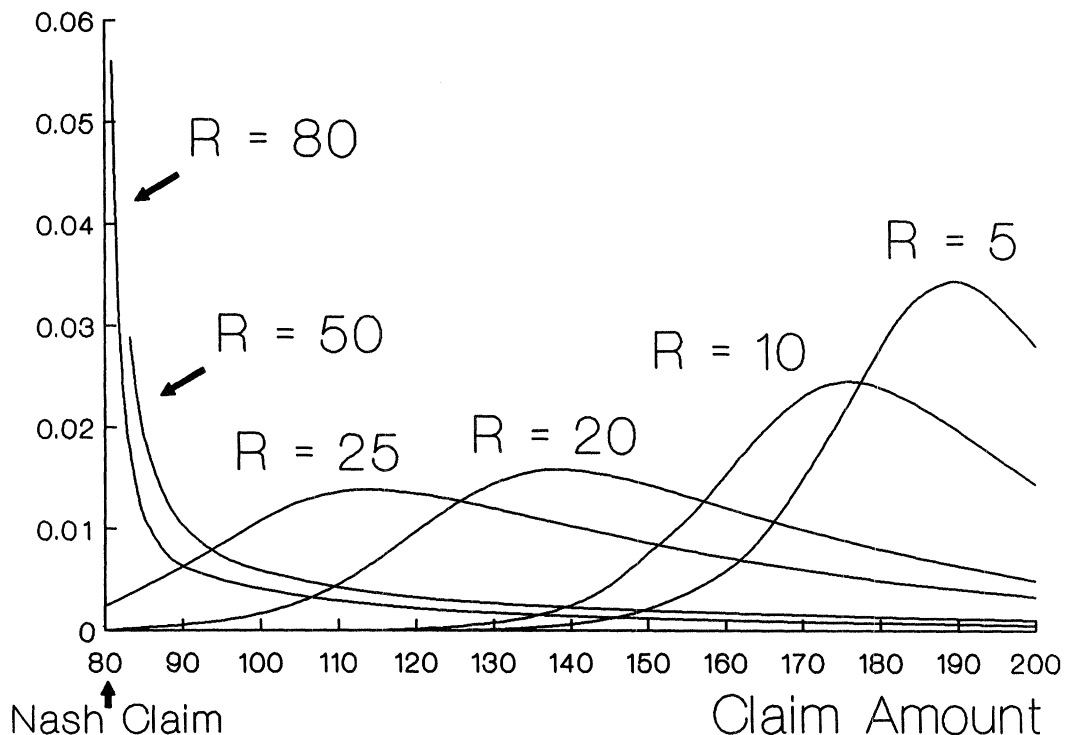
These plots reveal a clear inverse relationship between predicted claims and the magnitude of the penalty/reward parameter, as observed in the experiment. In particular, notice that the noise introduced in the logit equilibrium model does more than spread the predictions away from a central tendency at the Nash equilibrium. In fact, for low values of  $R$ , the claim distributions are centered well away from the Nash prediction, at the *opposite* end of the range of feasible choices.

We can use the logit equilibrium model (for  $\mu = 8.3$ ) to obtain predictions for the last three periods. These predictions, calculated from the equilibrium probability distributions in Figure 3, are found in the fifth row in Table 3. The closeness of the logit equilibrium predictions and the actual averages (row 2) for the final three periods is remarkable. In all cases, the predictions are much better than those of the Nash equilibrium, which is 80 for all treatments (row 6 in Table 3).<sup>14</sup> *To summarize, the estimated error parameter of the logit equilibrium model can be used to derive predicted average claims that track the salient treatment effect on claim data in the final three periods, an effect that is not explained by the Nash equilibrium.*

The logit-equilibrium approach in this section does not explain all aspects of the data. For example, the claims in part B are generally lower when preceded by very low claims in a competitive part A treatment, as can be seen in Figure 2. This cross-game learning, which has been observed in other experiments, is difficult to model, and is not surprising. After all, the optimal decision depends on beliefs about others' behavior, and low claims in a previous treatment can affect these beliefs. Beliefs would also be influenced by knowing the true price of the item that was lost in the traveler's dilemma game. This true value might be a focal point for claims made in early periods. Another aspect of the data that is not explained by the logit equilibrium model is the tendency for a significant

<sup>14</sup> We have no formal proof that the belief distributions in the logit learning model will converge to the equilibrium distributions, but notice that the simulated average claims in row 4 end up being reasonably close to the predicted equilibrium claims in row 5, even after as few as 8–10 simulated matchings, and the difference is largely due to the higher error parameter estimate for the dynamic model.

## Density

FIGURE 3. LOGIT EQUILIBRIUM CLAIM DENSITIES (WITH  $\mu = 8.3$ )

fraction of the subjects to use the same decision as in the previous period. Other subjects change their decisions frequently, even when average decisions have stabilized. There seems to be some inertia in decision-making that is not captured by the logit model. Finally, separate estimates of the logit error parameter for each treatment reveal some differences. However, the estimates are, with one exception, of the same order of magnitude and are similar to estimates that we have found for other games.

## VI. Conclusion

Basu's traveler's dilemma is of interest because the stark predictions of the unique Nash equilibrium are at odds with most economists' intuition about how people would behave in such a situation. This conflict between theory and intuition is especially sharp for low values of the penalty/reward parameter, since upward deviations from

the low Nash equilibrium claims are relatively costless. The experiment reported here is designed to exploit the invariance of the Nash prediction with respect to changes in the penalty/reward parameter. The behavior of financially motivated subjects confirmed our expectation that the Nash prediction would fail on two counts: claims were well above the Nash prediction for some treatments, and average claims were inversely related to the value of the penalty/reward parameter. Moreover, these results cannot be explained by any theory, static or dynamic, that is based on (perfectly rational) best responses to a previously observed claim, since the best response to a given claim is independent of the penalty/reward parameter in the traveler's dilemma game. In particular, the strong treatment effects are not predicted by learning direction theory, imitation theories, or evolutionary models that specify partial adjustments to best responses to the most recent outcome.

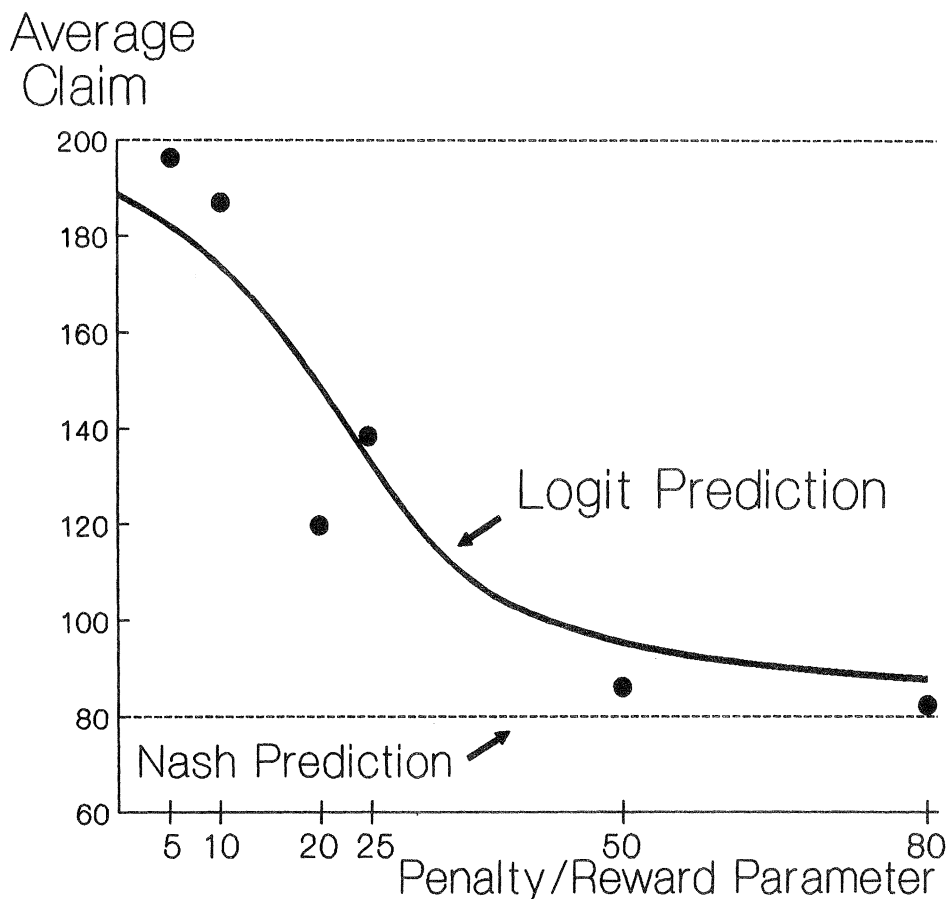


FIGURE 4. PREDICTED AND ACTUAL AVERAGE CLAIMS FOR THE FINAL THREE PERIODS

Note: Dots represent average claims for each of the treatments.

The Nash equilibrium is the central organizing concept in game theory, and has been for over 25 years. This approach should not be discarded; it has worked well in many contexts, and here it works well for high values of the penalty/reward parameter. Rather, what is needed is a generalization that includes the Nash equilibrium as a special case, and that can explain why it predicts well in some contexts and not others. One alternative approach is to model the formation of beliefs about others' decisions, and we implement this by estimating a dynamic learning model in which players make noisy best responses to beliefs that evolve, using a standard logit probabilistic

choice rule. In an equilibrium where beliefs stabilize, the belief and decision distributions are identical, although the probabilistic choice function will keep injecting some noise into the system.

The logit equilibrium model uses the logit probabilistic choice function to determine decisions, while keeping a Nash-like consistency-of-actions-and-beliefs condition. This model performs particularly well in the traveler's dilemma game, where the Nash predictions are at odds with both data and intuition about average claims and incentive effects. Consider the results for each treatment, as shown by the dark dots in Figure

4 that represent average claims for the final three periods plotted above the corresponding penalty/reward parameter on the horizontal axis. If one were to draw a freehand line through these dots, it would look approximately like the dark curved line, which is in fact the graph of the logit equilibrium prediction as a function of the  $R$  parameter (calculated on basis of the estimated value of the logit error parameter for the equilibrium model).<sup>15</sup> Even the treatment reversal between  $R$  values of 20 and 25 seems unsurprising given the closeness of these two treatments on the horizontal axis and the flatness of the densities for these treatments in Figure 3. Recall that the Nash prediction is 80 (the horizontal axis) for all treatments. Thus the data are concentrated at the *opposite* end of the range of feasible claims for low values of the penalty/reward parameter, which cannot be explained by adding errors around the Nash prediction. These data patterns are well explained by the logit equilibrium and learning models.

#### REFERENCES

- Anderson, Lisa and Holt, Charles A. "Information Cascades in the Laboratory." *American Economic Review*, December 1997, 87(5), pp. 847–62.
- Anderson, Simon P.; Goeree, Jacob K. and Holt, Charles A. "Rent Seeking with Bounded Rationality: An Analysis of the All-Pay Auction." *Journal of Political Economy*, August 1998a, 106(4), pp. 828–53.
- . "A Theoretical Analysis of Altruism and Decision Error in Public Goods Games." *Journal of Public Economics*, November 1998b, 70, pp. 297–323.
- . "Logit Equilibria for Auction-Like Games." Working paper, University of Virginia, 1998c.
- Basu, Kaushik. "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory." *American Economic Review*, May 1994 (*Papers and Proceedings*), 84(2), pp. 391–95.
- Brandts, Jordi and Holt, Charles A. "Naive Bayesian Learning and Adjustments to Equilibrium in Signaling Games." Working paper, University of Virginia, 1996.
- Camerer, Colin and Ho, Teck-Hua. "Experience Weighted Attraction Learning in Normal-Form Games." *Econometrica*, 1999 (forthcoming).
- Capra, C. Monica; Goeree, Jacob K.; Gomez, Rosario and Holt, Charles A. "Learning and Noisy Equilibrium Behavior in an Experimental Study of Imperfect Price Competition." Working paper, University of Virginia, 1998.
- Chen, Yan and Tang, Fang-Fang. "Learning and Incentive Compatible Mechanisms for Public Goods Provision: An Experimental Study." *Journal of Political Economy*, June 1998, 106(3), pp. 633–62.
- Cooper, David J.; Garvin, Susan and Kagel, John H. "Signalling and Adaptive Learning in an Entry Limit Pricing Game." *Rand Journal of Economics*, Winter 1997, 28(4), pp. 662–83.
- Crawford, Vincent P. "Adaptive Dynamics in Coordination Games." *Econometrica*, January 1995, 63(1), pp. 103–44.
- Erev, Ido and Roth, Alvin E. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." *American Economic Review*, September 1998, 88(4), pp. 848–81.
- Fudenberg, Drew and Levine, David K. *Learning in games*. Cambridge, Massachusetts: MIT Press, 1998.
- Lopez, Gladys. "Quantal Response Equilibria for Models of Price Competition." Ph.D. dissertation, University of Virginia, 1995.
- Luce, R. Duncan. *Individual choice behavior*. New York: Wiley, 1959.
- McKelvey, Richard D. and Palfrey, Thomas R. "Quantal Response Equilibria for Normal Form games." *Games and Economic Behavior*, July 1995, 10(1), pp. 6–38.
- . "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics*, 1998, 1(1), pp. 9–41.
- Mookherjee, Dilip and Sopher, Barry. "Learning and Decision Costs in Experimental Constant Sum Games." *Games and Economic*

<sup>15</sup> Since we estimated a virtually identical error rate for a different experiment in Capra et al. (1998), the predictions in Figure 4 could have been derived from a different data set.

- Behavior*, April 1997, 19(1), pp. 97–132.
- Ochs, Jack.** “Games with Unique Mixed Strategy Equilibria: An Experimental Study.” *Games and Economic Behavior*, July 1995, 10(1), pp. 202–17.
- Rhode, Paul and Stegeman, Mark.** “Non-Nash Equilibria of Darwinian Dynamics (with Applications to Duopoly).” Working paper, Virginia Polytechnical Institute, 1995.
- Rosenthal, Robert W.** “A Bounded-Rationality Approach to the Study of Noncooperative Games.” *International Journal of Game Theory*, 1989, 18(3), pp. 273–91.
- Roth, Alvin E. and Erev, Ido.** “Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term.” *Games and Economic Behavior*, January 1995, 8(1), pp. 164–212.
- Selten, Reinhard and Buchta, Joachim.** “Experimental Sealed Bid First Price Auctions with Directly Observed Bid Functions.” Working paper, University of Bonn, 1994.
- Smith, Vernon L. and Walker, James M.** “Monetary Rewards and Decision Cost in Experimental Economics: An Extension,” Working paper, University of Arizona, 1997.
- Vega-Redondo, Fernando.** “The Evolution of Walrasian Behavior.” *Econometrica*, March 1997, 65(2), pp. 375–84.