

OPTIMIZATION INCENTIVES AND COORDINATION
FAILURE IN LABORATORY STAG HUNT GAMES

BY RAYMOND BATTALIO, LARRY SAMUELSON, AND JOHN VAN HUYCK¹

1. INTRODUCTION

THE SPECIFICATION OF THE FEASIBLE strategies and preferences that define a strategic-form game, together with the assumption that players are substantively rational, provides a powerful framework for analyzing strategic behavior. This framework in turn can be summarized by the game's best-response correspondence. For example, one need only know the best-response correspondence of a strategic-form game to identify its Nash equilibria. The classical approach to games typically either exploits *only* the information contained in the best-response correspondence, or augments this information with risk-dominance and payoff-dominance considerations in order to choose between strict Nash equilibria.²

This paper reports an experimental investigation of three *stag hunt* games. The three games have identical best-response correspondences as well as similar payoff magnitudes, but produce different behavior.

Games $2R$, R , and $0.6R$, shown in Figures 1, 2, and 3, were used in the experiment. In each game, strategy X is a strict best response to any mixture that attaches a probability greater than q^* to X , where $q^* = 0.8$, while Y is a strict best-response to any mixture attaching a lower probability to X . Each game has two pure-strategy equilibria, where (X, X) is payoff dominant and (Y, Y) is risk dominant, as well as a mixed equilibrium in which X is played with probability q^* .

Our analysis of games $2R$, R , and $0.6R$ is motivated by the observation that the pecuniary incentive to select a best-response to an opponent's strategy is twice as large in game $2R$ as it is in game R and six tenths as large in game $0.6R$ as it is in game R . We call this incentive, given by the difference between the payoff of the best response to an opponent's strategy and the inferior response, the *optimization premium*. The optimization premium may be irrelevant to substantively rational agents, but we expect people to more readily learn to play a best response when the optimization premium is large, and expect the differing optimization premia of games $2R$, R , and $0.6R$ to induce systematically different play in laboratory experiments.

¹ We thank Menesh Patel, Bill Rankin, and Nick Rupp for research assistance, Simon Anderson, John Kagel, Jack Ochs, Richard McKelvey, and John Nachbar for helpful discussions, Dan Friedman, Robert Forsythe, Paul Straub, Martin Sefton, and their collaborators for making their data available to us, and two referees for helpful comments. Eric Battalio implemented the experimental design on the TAMU economic research laboratory network. The National Science Foundation and the Texas Advanced Research Program provided financial support. The first draft of this paper was called "Risk Dominance, Payoff Dominance, and Probabilistic Choice Learning," which was drafted while Van Huyck was on faculty development leave at the University of Pittsburgh.

² Hillas (1990) introduces a reformulation of Kohlberg and Mertens' (1986) strategic stability that makes the exclusive reliance on the best-response correspondence particularly obvious. Among theories that make an equilibrium selection in the stag hunt game, Carlsson and van Damme (1993) and Harsanyi (1995) choose the risk-dominant equilibrium, while Anderlini (1999) and Harsanyi and Selten (1988) choose the payoff-dominant equilibrium.

	X	Y
X	45,45	0,35
Y	35,0	40,40

FIGURE 1.—Game $2R$.

	X	Y
X	45,45	0,40
Y	40,0	20,20

FIGURE 2.—Game R .

	X	Y
X	45,45	0,42
Y	42,0	12,12

FIGURE 3.—Game $0.6R$.

To the extent possible, games $2R$, R , and $0.6R$ involve payoffs of similar magnitudes. In particular, the expected payoff from the mixed equilibrium is 36 for all three games. One can think of the optimization premium as describing the steepness, rather than the level, of the payoff function near an equilibrium. A larger optimization premium implies that the penalty for inferior play is larger.

Our experimental results provide evidence that changing the optimization premium influences behavior. The sensitivity of individual subjects to the history of opponents' play is greater in games with a larger optimization premium. Behavior converges more quickly in game $2R$ than in R , and more quickly in game R than in game $0.6R$. The payoff-dominant equilibrium is more likely to emerge the smaller is the optimization premium.

2. EXPERIMENTAL DESIGN

The experiment consists of three treatments. Each treatment consists of eight cohorts. Eight subjects participated in each cohort. Each cohort plays one of the three games, either $2R$, R , or $0.6R$, seventy-five times. We used a single-population random matching protocol to pair subjects within a cohort. The subjects were informed that they were being randomly paired.

The subjects had common and complete information about both their own and everybody else's earnings table. Actions were labeled 1 and 2, and each subject chose one such action in each period. After their choices were made, the subjects were randomly paired with an anonymous opponent to determine an outcome for each pair. Since outcomes were reported privately, subjects could not use common information about the outcomes in previous periods to coordinate on an equilibrium.

Cell entries in Figures 1, 2, and 3 denote the number of cents earned by a subject pair for each action combination in each round. Earnings were presented in matrix form and subjects were instructed on how to derive the other participant's earnings from the earnings table.

No preplay communication was allowed. Messages were sent electronically on a PC-network.

The subjects were recruited from undergraduate economics classes at Texas A&M University in the Spring of 1996, Fall of 1997, and Spring of 1998. A total of 192 subjects participated in the experiment: eight cohorts of eight subjects in three treatments. After reading the instructions, but before the session began, the subjects filled out a questionnaire to determine that they understood how to read earnings tables.³ A session lasted about two hours. Repeated play of the payoff-dominant equilibrium for seventy-five periods results in a subject earning \$33.75.

3. OPTIMIZATION INCENTIVES

Games $2R$, R , and $0.6R$ differ in the penalty attached to not playing a best-response or, more optimistically, in the premium for playing a best-response. We refer to this incentive as the *optimization premium*. Let $\pi_j(X, q)$ denote the expected payoff to a player in game j who plays X and expects his opponent to play X with probability q . Let $\pi_j(Y, q)$ be similarly defined for Y . Then the optimization premium for game j is the function $r_j(q): [0, 1] \rightarrow \mathbb{R}$ given by

$$\begin{aligned} r_{2R}(q) &= \pi_{2R}(X, q) - \pi_{2R}(Y, q) = 50(q - q^*) = \delta_{2R}(q - q^*), \\ r_R(q) &= \pi_R(X, q) - \pi_R(Y, q) = 25(q - q^*) = \delta_R(q - q^*), \\ r_{0.6R}(q) &= \pi_{0.6R}(X, q) - \pi_{0.6R}(Y, q) = 15(q - q^*) = \delta_{0.6R}(q - q^*), \end{aligned}$$

where δ_j is the *optimization premium parameter*. Hence, for any opponent's strategy q , the optimization premium is twice as large in game $2R$ as it is in game R and six tenths as large in game $0.6R$ as it is in game R .

Our intuition is that the process attracting players to choose best-responses will be more effective in games in which the optimization premium is larger. To make this precise, consider the following probabilistic choice model that can be derived axiomatically (see Luce (1959)) or from a random utility framework (see Maddalla (1983) and Anderson, de Palma, and Thisse (1992)):

$$p(q, \lambda, j) = \frac{\exp(\lambda \pi_j(X, q))}{\exp(\lambda \pi_j(X, q)) + \exp(\lambda \pi_j(Y, q))},$$

where $p(q, \lambda, j)$ is the probability that X is chosen, given q and λ , in game j , and λ is a precision parameter. We can solve for the logistic-response function

$$p(q, \lambda, \delta_j) = \frac{\exp(\lambda \delta_j (q - q^*))}{1 + \exp(\lambda \delta_j (q - q^*))}.$$

If λ equals 0, players mix equally over all strategies, while λ sufficiently large gives essentially best-response behavior. Holding λ constant, subjects' behavior will be more responsive to q in game $2R$ than in game R and in game R than in game $0.6R$, since a

³ The instructions for the experiment are available on the web at "erl.tamu.edu" or "www.ssc.wisc.edu/~larrysam".

larger optimization parameter δ_j gives a logistic-response function closer to the best-response function:⁴

HYPOTHESIS 1: Subjects' behavior will be more responsive to beliefs the larger is the optimization premium parameter.

Following Fudenberg and Levine (1998), we can use the logistic-response function to define a single-population continuous-time logistic-response dynamic,

$$\dot{q} = p(q, \lambda; \delta_j) - q,$$

where q is reinterpreted as the frequency of action X in the population and it is assumed that the population is sufficiently large as to allow the random individual choices to be captured by a deterministic population equation.⁵

Figure 4 illustrates this dynamic for the case of $\lambda = 1$. For any finite $\lambda > 0$, the magnitude of the change in the population state q , and hence the speed of convergence, differs by optimization premia.

HYPOTHESIS 2: Behavior will converge to an equilibrium more quickly the larger is the optimization premium.

This result is typical of noisy belief-based models in which players react more vigorously to beliefs when payoff differences are larger. Common models of population behavior based on deterministic or stochastic generalizations of the replicator dynamic similarly assume that rates of adjustment are increasing in the current difference in payoffs between strategies (for example, Binmore, Gale, and Samuelson (1995), Borgers and Sarin (1997), or Weibull (1995)).

Fixing λ , a *logit equilibrium* is a fixed point of the two players' logistic-response functions (McKelvey and Palfrey (1995)). The stationary states of the single-population logistic-response dynamic correspond to symmetric logit equilibria.

Figure 4 graphs the logistic-response dynamic for the case of $\lambda = 1$. For comparison, it also graphs the single-population continuous-time best-response dynamic, which is the same for all three games. Games $2R$ and R have three logit equilibria that are close to the best-response equilibria, with the "risk-dominant" equilibrium having a larger basin of attraction in the case of game R than game $2R$, and with both basins of attraction being larger than in the case of the best-response dynamic.⁶ Game $0.6R$ has a single logit equilibrium (given $\lambda = 1$), which is close to the risk-dominant equilibrium, and whose basin of attraction comprises the entire state space.

⁴ A growing literature examines models of behavior in games. Rather than a complete model of adaptive behavior, our goal is to answer the question, "Does the optimization premium matter?", which is most effectively answered within the context of the logit response function.

⁵ Crawford (1995) examines an alternative belief-based dynamic. Borgers and Sarin (1997), Binmore and Samuelson (1997), Binmore, Gale, and Samuelson (1995), Erev and Roth (1998), and Roth and Erev (1995) model current actions as functions of previous experience, with favorable experiences reinforcing the tendency to take an action. In practice, beliefs are typically estimated as a function of previous outcomes, bringing the two types of model closer together (see Hopkins (1999)). More general models include Camerer and Ho's (1999) experience-weighted attraction model and Stahl's (1996, 1999) rule-learning models.

⁶ This observation is a consequence of the way the logit equilibrium close to the mixed equilibrium changes as players become imprecise in their responses (Fudenberg and Levine (1998)).

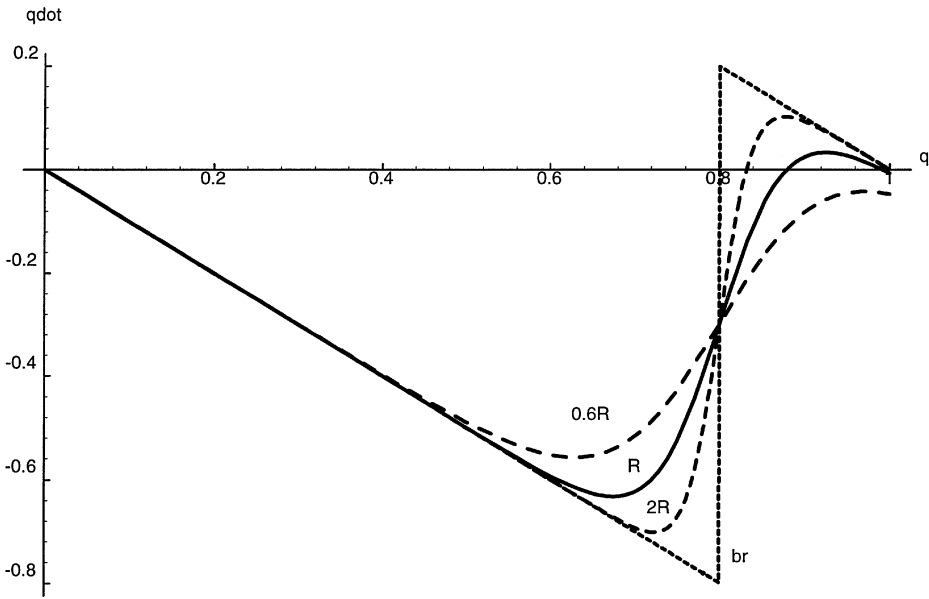


FIGURE 4.—One population continuous time best-response and logistic-response dynamics ($\lambda = 1$).

For any finite $\lambda > 0$, the basin of attraction of the logit equilibrium closest to the risk dominant equilibrium expands as the optimization premium falls, until a sufficiently low optimization premium is reached that there is a single logit equilibrium, closer to the risk-dominant than the payoff-dominant equilibrium. If we think of some fixed distribution governing the initial condition of the dynamic, then the effect of probabilistic choice is to make the payoff-dominant equilibrium less likely than in the case of best-response dynamics, and less likely as the optimization premium is smaller.

This result is somewhat counterintuitive. Learning is likely to be noisy. We would expect a smaller optimization premium to increase the likelihood that noisy learning induces the population to enter the basin of attraction of the payoff-dominant equilibrium (X, X) .⁷ A variety of forces may be behind this result, one of which is captured by the aspiration-and-imitation model of Binmore and Samuelson (1997). In their model, players are more likely to revise their strategies whenever their payoffs fall below an aspiration level. Learning is thus noisier when payoffs are smaller, and the population is more likely to stumble away from the neighborhood of an equilibrium if the latter involves relatively low payoffs. Hence, whenever the risk-dominant and payoff-dominant equilibria differ, the learning process is more likely to cause the proportion of the population playing strategy X to move away from the relatively low-payoff risk-dominant equilibrium than from the payoff-dominant equilibrium, and this difference is more

⁷ When the optimization premium is smaller, we expect considerations other than expected-payoff calculations to become more important in shaping behavior. Analysis is likely to give way to behavioral rules and payoff consequences are likely to be assessed not by calculation but by experimentation, in the form of simply playing a strategy to see what happens. Learning thus becomes noisier.

CONTINGENCY TABLE I
TREATMENT BY PERIOD 1 SUBJECT CHOICE

	<i>X</i>	<i>Y</i>	Total
0.6 <i>R</i>	41 (0.64)	23 (0.36)	64 (1.00)
<i>R</i>	45 (0.70)	19 (0.30)	64 (1.00)
2 <i>R</i>	34 (0.53)	30 (0.47)	64 (1.00)
Total	120 (0.63)	72 (0.37)	192 (1.00)

CONTINGENCY TABLE II
TREATMENT BY PERIOD 75 SUBJECT CHOICE

	<i>X</i>	<i>Y</i>	Total
0.6 <i>R</i>	28 (0.44)	36 (0.56)	64 (1.00)
<i>R</i>	16 (0.25)	48 (0.75)	64 (1.00)
2 <i>R</i>	3 (0.05)	61 (0.95)	64 (1.00)
Total	47 (0.24)	145 (0.76)	192 (1.00)

pronounced the smaller is the optimization premium.⁸ This leads to a prediction that is not made by best-response, logistic-response, or replicator dynamics:

HYPOTHESIS 3: Behavior is more likely to converge to the payoff-dominant equilibrium the smaller is the optimization premium.

4. EXPERIMENTAL RESULTS

4.1. Treatment Behavior

In period 1, 63 percent of the subjects play *X*, the payoff-dominant action. Risk dominance is thus not a salient deductive selection principle, though not enough subjects focus on payoff dominance to make playing the payoff-dominant action a best-response, since 0.63 is less than q^* .

Contingency Table I, crossing treatment, and subject choice in period 1, can be used to test the hypothesis that initial behavior did not vary by treatment. The Chi-square statistic is 4.1 which, given 2 degrees of freedom, has a *p-value* of 0.13. Hence, subjects' slight tendency to initially play the payoff-dominant action more frequently when the optimization premium is smaller is not statistically significant at conventional levels.

The insignificant difference in initial behavior across treatments grows to a large treatment effect by the end of the session. Contingency Table II shows that in period 75, only 5 percent of subjects in treatment 2*R* play action *X*, while 44 percent of subjects in

⁸ Similar considerations appear in the heterogeneous-payoff model of Myatt and Wallace (1997). In contrast, Kandori, Mailath, and Rob (1993) and Young (1993) use evolutionary arguments based on the best-response function to select the risk-dominant equilibrium of a stag hunt game, regardless of the optimization premium, while Robson and Vega-Redondo (1996) use a similar model to select the payoff-dominant equilibrium. Friedman (1996) suggests that a population may be more likely to move away from the risk-dominant equilibrium as a result of subjects' efforts to "teach" others that the payoff-dominant equilibrium would be better, though this intuition contrasts with the theoretical results of Ellison (1997). See also Camerer, Ho, and Chong (2000), and compare Van Huyck, Cook, and Battalio (1997).

TABLE III
THE AVERAGE CHANGE IN x GIVEN x

Treatment	0	1	2	3	4	5	6	7	8
0.6R	0.55	0.38	0.10	-0.078	-0.27	-0.03	-0.34	-0.20	-0.16
R	0.23	-0.18	-0.22	0.14	-0.40	-0.37	-0.15	0.19	-0.32
2R	0.07	-0.26	-0.46	0.16	-0.50	-0.57	-0.38	-1.75	

treatment 0.6R are still playing action X . The payoff-dominant action is thus more prevalent in games with smaller optimization premia.

To gain some insight into the dynamics behind these outcomes, let state x denote the number of subjects choosing action X in a cohort in a period. It ranges from 0 to 8. Table III reports the average of the change in x , denoted by Δx , for each state and treatment. For every x in the interval $\{2, 3, 4, 5\}$, larger optimization premia are associated with average changes whose absolute values are larger, though two of the changes appear to go in the wrong direction in the case of state 3. In contrast, for states near the risk-dominant equilibrium, the largest average changes are attached to the 0.6R treatment, which exhibits a strong tendency to move away from the risk-dominant equilibrium. This suggests that something beyond the considerations captured by the logistic choice model, such as an aspiration-based desire to avoid exceptionally low payoffs, is at work, pushing the population toward the payoff-dominant equilibrium when the optimization premium is small.

Figure 5 supplements Table III by reporting the count for each value of Δx that goes into the average change in x . The figure is truncated at ± 4 , because no value of x ever changed by more than ± 3 from one period to the next. Figure 5 shows that no value of x is perfectly absorbing. However, in treatment 0.6R, the state with the largest count for $\Delta x = 0$ was state $x = 8$, the payoff-dominant equilibrium, while for the other two games the largest count for $\Delta x = 0$ was at state $x = 0$, the risk-dominant equilibrium. This pattern remains if we normalize the counts by dividing through by the number of times each state x arose in a treatment.⁹

4.2. Cohort Behavior

Our analysis of the results by treatment suggests that initial behavior varies little across treatments, but experience teaches subjects to play the risk-dominant action more effectively the larger the optimization premium. In this section, we examine the data by cohort to develop an understanding of how this happens.

Table IV reports the initial and terminal outcome by cohort. All 24 of the cohorts start in the basin of attraction of the risk-dominant equilibrium (Y, Y). Three 0.6R cohorts, four R cohorts, and five 2R cohorts implement an equilibrium in period 75. This observation is consistent with hypothesis 2: cohorts with a larger optimization premium were more likely to have converged to an equilibrium by the end of the session.

⁹ If we examine states near the best-response separatrix in Figure 5, that is, states $x = 6$ and $x = 7$, we do not find that movements toward the payoff dominant equilibrium (upward) are especially likely when the optimization premium is small (compare the 0.6R and R cases). Because our games have identical mixed-equilibrium payoffs, differences in the behavior predictions of the aspiration and imitation model, across optimization premia, disappear as the population approaches the separatrix.

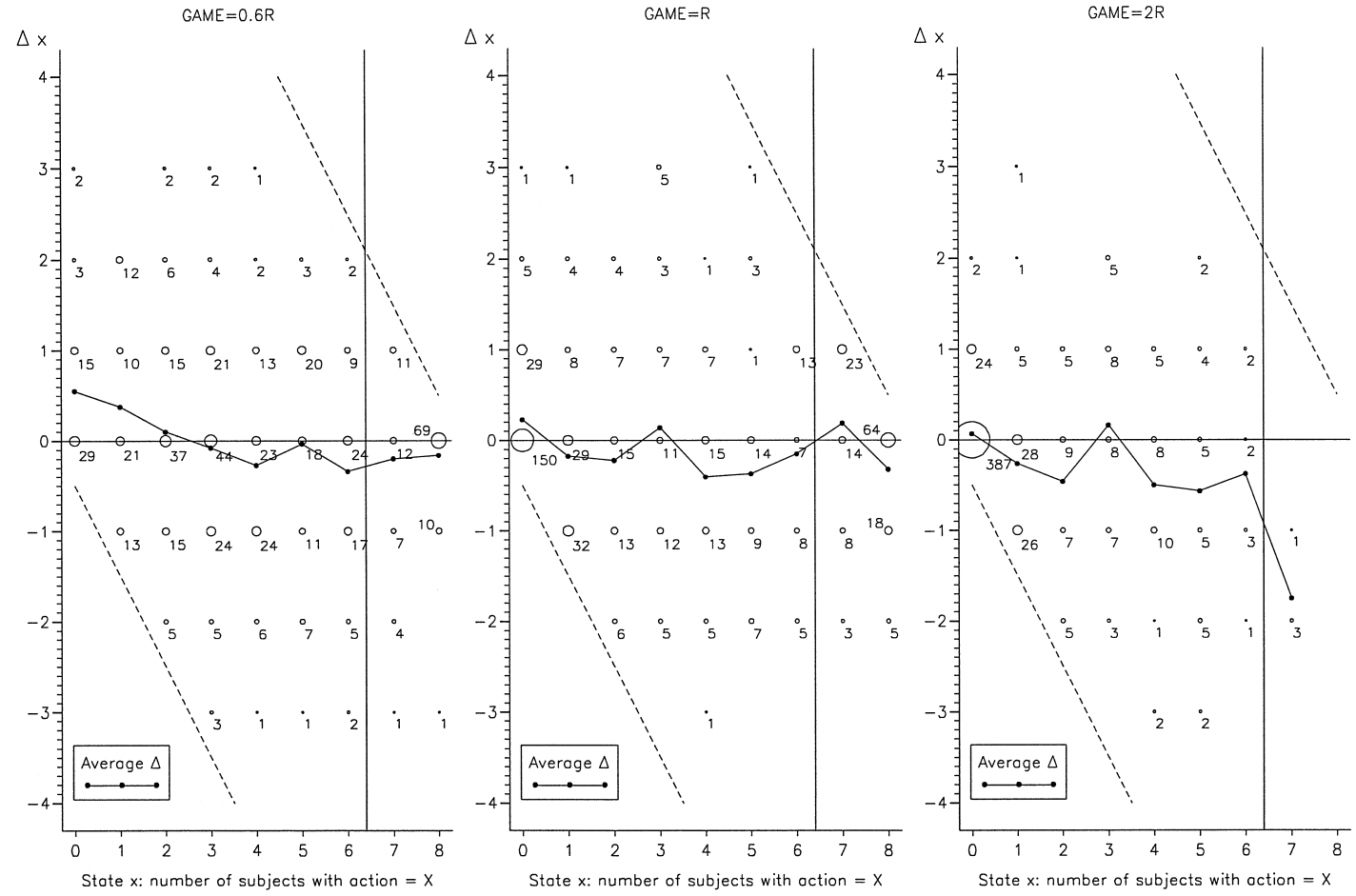


FIGURE 5.—Change in x , Δx , as a function of x all periods by treatment.

TABLE IV
RANKING OVERALL FREQUENCY OF X BY COHORT

Game	Cohort	Initial Outcome	Terminal Outcome	Overall Frequency of X	Rank	Average Per Capita Earnings
0.6R	1	0.500	0.375	0.540	20	\$19.67
0.6R	2	0.750	1.000 ^a	0.803	21	\$27.35
0.6R	3	0.750	1.000 ^a	0.930	23	\$31.31
0.6R	4	0.750	0.250	0.397	17	\$16.31
0.6R	5	0.750	0.125	0.508	19	\$19.39
0.6R	6	0.500	0.375	0.218	13	\$12.59
0.6R	7	0.625	0.375	0.343	16	\$14.76
0.6R	8	0.500	0.000	0.213	12	\$12.86
R	9	0.750	0.125	0.145	7	\$15.88
R	10	0.625	0.000	0.203	11	\$16.12
R	11	0.750	0.625	0.920	22	\$31.00
R	12	0.750	0.000	0.255	15	\$17.06
R	13	0.750	0.125	0.418	18	\$19.44
R	14	0.500	0.000	0.160	9	\$15.63
R	15	0.750	1.000 ^a	0.937	24	\$31.44
R	16	0.750	0.125	0.098	5	\$15.38
2R	17	0.625	0.125	0.242	14	\$27.22
2R	18	0.500	0.000	0.048	3	\$28.99
2R	19	0.375	0.000	0.017	1	\$29.44
2R	20	0.625	0.000	0.152	8	\$27.01
2R	21	0.375	0.000	0.107	6	\$28.03
2R	22	0.500	0.125	0.035	2	\$29.19
2R	23	0.750	0.000	0.055	4	\$28.51
2R	24	0.500	0.125	0.185	10	\$26.63

^aSeparatrix crossing between initial and terminal outcome.

Which equilibrium emerges? Table IV indicates that, by the terminal outcome, one of the 0.6R cohorts (8), three of the R cohorts (10, 12, and 14), and five of the 2R cohorts (18, 19, 20, 21, and 23) converged to the risk-dominant equilibrium. Conversely, two 0.6R cohorts (2 and 3 respectively) and one R cohort (15) converged to the payoff-dominant equilibrium, while none of the 2R cohorts converged to the payoff-dominant equilibrium. Notice that, given the observed initial conditions, a cohort must cross the (best-response) separatrix to converge to the payoff-dominant equilibrium. Cohorts were less likely to escape from the risk dominant equilibrium's best-response basin of attraction the larger the optimization premium. Our results are thus consistent with Hypothesis 3: the payoff-dominant equilibrium emerged less frequently in treatments with a larger optimization premium.

A nonparametric rank sum test reveals that the observed difference in behavior is statistically significant. In particular, Table IV ranks the cohorts by the overall frequency of the payoff-dominant action X . The cohort with the lowest X frequency is 19 and it receives a rank of 1. The cohort with the highest X frequency is 15 and it receives a rank of 24. A quick inspection of the rankings reveal that the 2R cohorts tend to receive single digit rankings and the 0.6R cohorts all receive double digit rankings.

These rankings can be used to perform the Kruskal-Wallis multiple comparisons test, which is based on the sum of the ranks by treatment (Conover (1980, p. 231)). The rank sum for the 0.6R cohorts is 48, for the R cohorts is 111, and for the 2R cohorts is 141.

The null hypothesis is no treatment difference. The test statistic is 11.3, which is approximately Chi-square. The probability value of 0.0036 rejects the null hypothesis under all conventional levels of statistical significance.

Given the alternative hypothesis of treatment differences, we proceed to determine which pairs of treatments differed. Dividing the rank sum by the number of observations, 8, gives the normalized rank sum. The absolute value of the difference in the normalized rank sum between the $0.6R$ and R treatments is 7.9 and between $0.6R$ and $2R$ is 11.6. Both of these values exceed the critical value of 7.5 at the 1 percent level of statistical significance. Hence, we conclude behavior in the $0.6R$ treatment was different than in the other two treatments. The absolute value of the normalized rank sum difference between the R and $2R$ treatments is 3.75, which is not statistically significant.

Figure 6 reports the five-period mean frequency of the payoff-dominant action by cohort. The three horizontal reference lines denote the frequencies with which X is played in the risk-dominant equilibrium (0.0), the mixed equilibrium (0.8), and the payoff-dominant equilibrium (1.0). The figure illustrates two results already derived from Table IV: Cohorts with a larger optimization premium were more likely to have converged to an equilibrium by the end of the session, and the payoff-dominant equilibrium emerged less frequently in treatments with a larger optimization premium.

As seen in Figure 6, it takes a long time to converge to a mutually consistent outcome. Amongst cohorts that converged to the risk-dominant equilibrium, it takes longer for R cohorts to reach the risk-dominant equilibrium than it does for $2R$ cohorts. If we examine the first (five period) state in which every subject in a cohort plays the risk-dominant action (excluding the R cohorts that never converge to the risk-dominant equilibrium), we find that the remaining six R cohorts take an average of 50 periods for all subjects to reach the risk-dominant equilibrium, while the eight $2R$ cohorts take an average of 26 periods. (Only one $0.6R$ cohort converges to the risk-dominant equilibrium state. It took 58 periods.) The evidence is thus consistent with the hypothesis that reducing the optimization premium reduces the speed of convergence to the inefficient risk-dominant equilibrium.

The results reported in Figure 6 reflect the qualitative features of our last two hypotheses. Convergence is more rapid when the optimization premium is larger. The risk-dominant equilibrium emerges as the customary way to play in all of the $2R$ cohorts and in six out of eight R cohorts. The risk-dominant equilibrium emerged only once in the eight $0.6R$ cohorts. Conversely, the payoff-dominant equilibrium emerges as the customary way to play in two $0.6R$ cohorts and one R cohort.

The last column of Table IV, reporting average per capita earnings by cohort, provides insight into the economic significance of these findings. The average subject in cohort 15 earned \$31.44, which was the highest average. Cohort 3 is a close second, earning \$31.31. The average subject in cohort 6 earned \$12.59, which was the least. A failure to coordinate on the payoff-dominant equilibrium was thus very costly to subjects in the $0.6R$ treatments in both absolute terms, about a \$19 difference, and in percentage terms, as cohort 6 earns only 40 percent of cohort 3's earnings. In contrast, cohorts in the $2R$ treatments lost much less as a result of the observed coordination failures.

4.3. Individual Behavior

Our examination of individual behavior begins with an estimation of the relationship between subjects' strategy choices and their experience. We first suppose the probability

that subject i attaches to her opponent playing strategy X , at time t , denoted by q_{it} , is given by

$$q_{it} = \frac{q_0 d^{t-1} + I_{i1} d^{t-2} + \dots + I_{i,t-2} d + I_{i,t-1}}{d^{t-1} + d^{t-2} + \dots + 1}$$

where q_0 is the prior probability, I_{it} equals one if i 's opponent played X at time τ and zero otherwise, and d is the discount factor. If $d = 1$, then this model yields fictitious play beliefs, and if $d = 0$, then we have Cournot beliefs. If we remove the prior, we get Cheung and Friedman's (1997) formulation.

We assume that the probability that subject i chooses strategy X at time t , denoted by p_{it} , is given by

$$p_{it} = \frac{\exp(\alpha_j + \beta_j(q_{it} - q^*))}{1 + \exp(\alpha_j + \beta_j(q_{it} - q^*))},$$

where $j \in \{2R, R, 0.6R\}$ indexes the games. When $\alpha_j = 0$ and $\beta_j = \lambda \delta_j$, this is the logistic response function discussed above, where λ is the precision parameter and δ_j is the optimization premium parameter. The constant term α_j is included to capture a possible tendency to move away from low payoffs, suggested by the data in Table III. Maximum likelihood estimates, computed using Gauss, are shown in Table V.

The estimated β have the expected ordering and are statistically different from each other: $\beta_{0.6R} < \beta_R < \beta_{2R}$. Individual subjects are more sensitive to the history of opponents' in games with a larger optimization premium.¹⁰

Because $\beta_j = \lambda \delta_j$, the estimated β_j reveal that the precision parameter λ is not constant across treatments. If it were, then not only would $\beta_{0.6R} < \beta_R < \beta_{2R}$, but $\beta_{0.6R}$ would equal $0.6\beta_R$, or 3.59, and β_{2R} would equal $2\beta_R$, or 11.96. Both $\beta_{0.6R}$ and β_R are more than two standard errors away from the estimated values, allowing us to reject the hypothesis of a stable precision parameter across treatments. The sensitivity of actions to the optimization premium appears to exhibit decreasing returns.

The estimated priors are remarkably close to the observed frequencies. The memory discount parameter estimates are plausible, but are closer to fictitious play than we expected.

The constant α_j is significantly positive in all three treatments, indicating a bias in favor of the payoff-dominant action. The bias does not appear to vary systematically with the payoff tables.¹¹ The logistic response model cannot accommodate the observed bias, as the logistic response function forces players to be indifferent between actions X and Y whenever they attach a probability of 0.8 to their opponent's playing X (see Figure 4), while the experimental subjects are significantly more likely to play X under such

¹⁰ We obtain analogous results if we impose the restriction that $\alpha_j = 0$ and hence work with the logistic response function.

¹¹ The aspiration and imitation model lying behind Hypothesis 3 suggests that strategy choices should be noisier when payoffs are small, and hence should be noisier near the payoff-dominant than the risk dominant equilibrium, with this difference most pronounced when the optimization premium is small. Translating these differing noise levels into differences in the trend α_j would require a richer specification, capturing such features of the aspiration and imitation model as the importance of an agent's previous choice and the current population state, though the ability of noisy choice to translate into an increased likelihood of absorption at the payoff dominant equilibrium suggests that α_j may be positive and decreasing in the optimization premium.

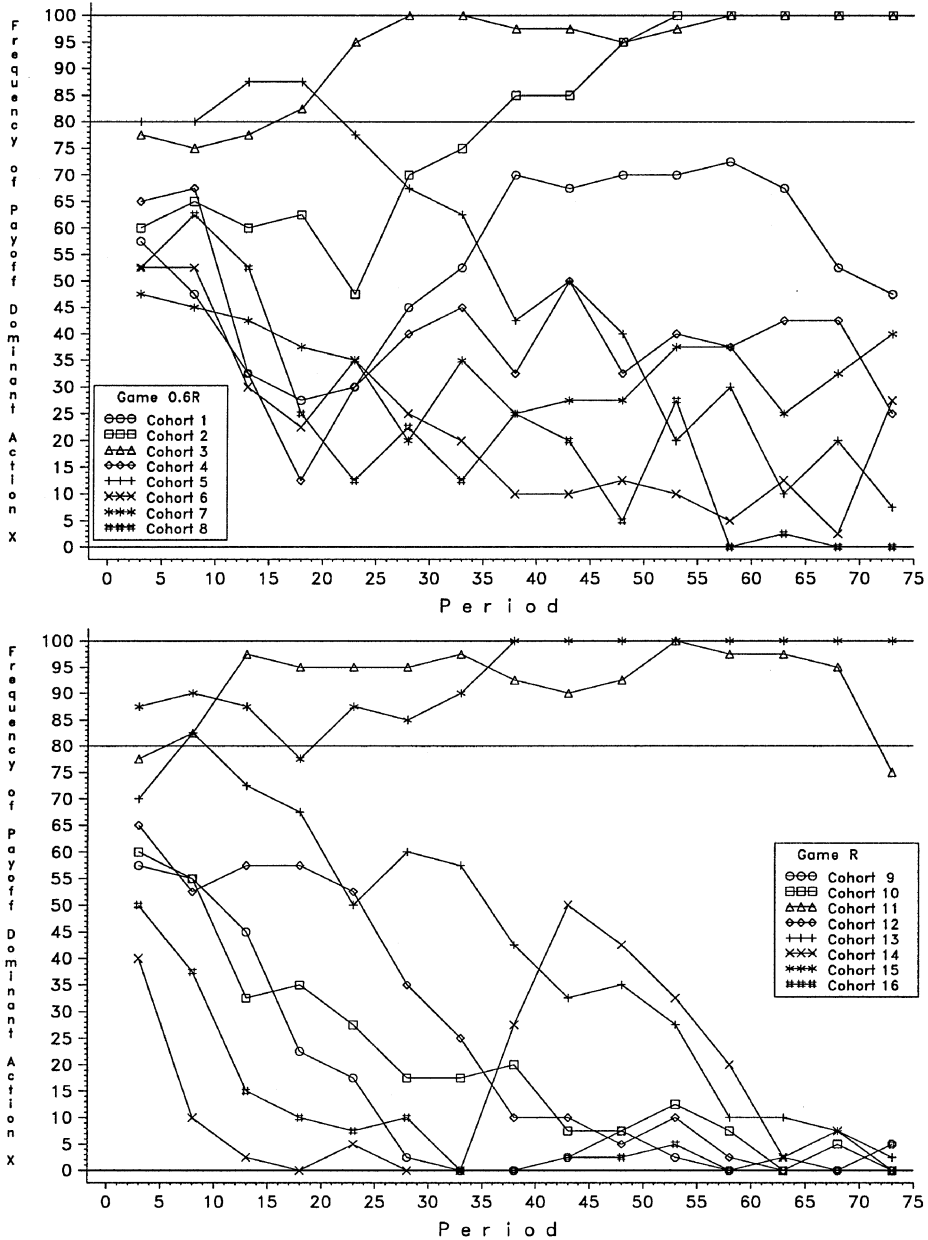


FIGURE 6.—Five period mean frequency of X by cohort.

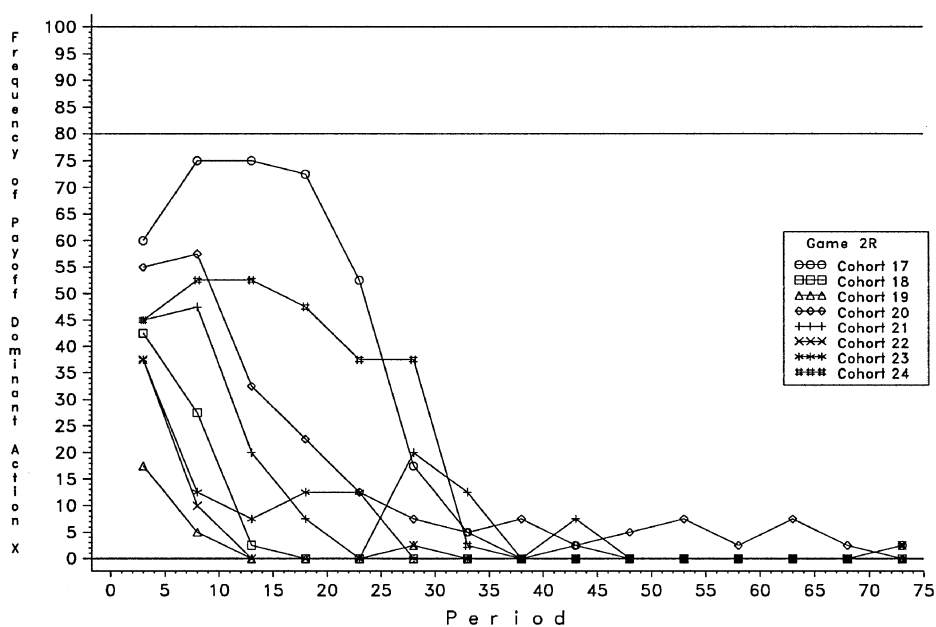


FIGURE 6.—Continued

circumstances. Our individual results are thus consistent with the first two hypotheses generated by the logistic choice model, but also suggest that the model fails to capture important aspects of observed behavior. More theoretical and experimental work is required to assess whether individual behavior matches the predictions of the aspiration and imitation model.

4.4. Literature Discussion

Experiments involving sequences of stag hunt games have been conducted by Clark, Kay, and Sefton (1996), Cooper, DeJong, Forsythe, and Ross (1992), Friedman (1996), Schmidt, Shupp, Walker, and Ostrom (1997), and Straub (1995). As is the case with our

TABLE V
ESTIMATED LOGISTIC RESPONSE MODEL
MEAN LOG-LIKELIHOOD -0.3637 ; NUMBER OF CASES 14400;
STANDARD ERRORS IN PARENTHESES

Treatment j	q_{0j}	d_j	α_j	β_j
0.6R	0.64 (0.05)	0.85 (0.01)	1.30 (0.06)	4.49 (0.15)
R	0.68 (0.04)	0.84 (0.02)	1.50 (0.07)	5.98 (0.16)
2R	0.65 (0.04)	0.89 (0.01)	1.28 (0.12)	6.95 (0.27)

results, play typically converges to the equilibrium whose best-response basin of attraction contains the initial outcome, with this equilibrium more likely to be the risk-dominant equilibrium the larger is the latter's basin of attraction. However, Schmidt, et al. observe three cases in which play begins in the risk-dominant basin of attraction but crosses the separatrix to converge to the payoff-dominant equilibrium.

Rankin, Van Huyck, and Battalio (1999) report an experiment in which subjects play a sequence of similar games in which payoffs, action labels, and game forms are constantly changing, forcing subjects to focus on abstract similarities between games. Payoff dominance emerges as an equilibrium selection principle even when the risk-dominant equilibrium has an extremely large basin of attraction, with values of q^* as large as 0.97.

5. CONCLUSION

Our results provide evidence that more than the best-response correspondence matters when predicting human behavior in laboratory experiments. We have focused on the optimization premium—the expected earnings difference between the two actions—in three stag hunt games that have the same best-response correspondence, the same mixed strategy equilibrium, and the same expected payoff at this mixed strategy equilibrium, but different pecuniary incentives to play a best-response.

We find statistically and economically significant evidence that the optimization premium helps explain observed behavior. The sensitivity of individual subjects to the history of opponents' play is greater in games with a larger optimization premium. Behavior converges more quickly the larger the optimization premium. The risk-dominant equilibrium is more likely to emerge the larger is the optimization premium.

*Dept. of Economics, Texas A&M University, College Station, TX 77843, U.S.A.,
Dept. of Economics, University of Wisconsin-Madison, 1180 Observatory Dr., Madison,
WI 53706, U.S.A.,*

and

Dept. of Economics, Texas A&M University, College Station, TX 77843, U.S.A.

Manuscript received January, 1998; final revision received June, 2000.

REFERENCES

- ANDERLINI, LUCA (1999): "Communication, Computability, and Common Interest Games," *Games and Economic Behavior*, 27, 1–37.
- ANDERSON, SIMON P., ANDRÉ DE PALMA, AND JACQUES-FRANÇOIS THISSE (1992): *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: The MIT Press.
- BINMORE, KEN, JOHN GALE, AND LARRY SAMUELSON (1995): "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior*, 8, 56–90.
- BINMORE, KEN, AND LARRY SAMUELSON (1997): "Muddling Through: Noisy Equilibrium Selection," *Journal of Economic Theory*, 74, 235–265.
- BORGERS, TILMAN, AND RAJIV SARIN (1997): "Learning through Reinforcement and the Replicator Dynamics," *Journal of Economic Theory*, 77, 1–14.
- CAMERER, COLIN, AND TECK-HUA HO (1999): "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica*, 67, 827–874.
- CAMERER, COLIN, TECK-HUA HO, AND JUIN-KUAN CHONG (2000): "Sophisticated EWA Learning and Strategic Teaching in Repeated Games," laser-script.

- CARLSSON, HANS, AND ERIC VAN DAMME (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.
- CHEUNG, YIN-WONG, AND DANIEL FRIEDMAN (1997): "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 10, 46–76.
- CLARK, KENNETH, STEPHEN KAY, AND MARTIN SEFTON (1996): "When Are Nash Equilibria Self-Enforcing? An Experimental Analysis," laser-script.
- CONOVER, W. J. (1980): *Practical Nonparametric Statistics*, 2nd Edition. New York, NY: John Wiley & Sons.
- COOPER, RUSSELL, DOUGLAS V. DEJONG, ROBERT FORSYTHE, AND THOMAS W. ROSS (1992): "Communication in Coordination Games," *Quarterly Journal of Economics*, 107, 739–773.
- CRAWFORD, VINCENT P. (1995): "Adaptive Dynamics in Coordination Games," *Econometrica*, 63, 103–144.
- EREV, IDO, AND ALVIN E. ROTH (1998): "On the Need for Low Rationality, Cognitive Game Theory: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88, 848–881.
- ELLISON, GLEN (1997): "Learning from Personal Experience: One Rational Guy and the Justification of Myopia," *Games and Economic Behavior*, 19, 180–210.
- FRIEDMAN, DANIEL (1996): "Equilibrium in Evolutionary Games: Some Experimental Results," *Economic Journal*, 106, 1–25.
- FUDENBERG, DREW, AND DAVID K. LEVINE (1998): *The Theory of Learning in Games*. Cambridge, MA: The MIT Press.
- HARSANYI, JOHN C. (1995): "A New Theory of Equilibrium Selection for Games with Complete Information," *Games and Economic Behavior*, 8, 91–122.
- HARSANYI, JOHN C., AND REINHARD SELTEN (1988): *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: The MIT Press.
- HILLAS, JOHN (1990): "On the Definition of the Strategic Stability of Equilibria," *Econometrica*, 58, 1365–1390.
- HOPKINS, ED (1999): "Two Competing Models of How People Learn in Games," laser-script, University of Pittsburgh and University of Edinburgh.
- KANDORI, MICHIIHIRO, GEORGE J. MAILATH, AND RAFAEL ROB (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61, 29–56.
- KOHLBERG, ELON, AND JEAN-FRANCOIS MERTENS (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003–1038.
- LUCE, R. DUNCAN (1959): *Individual Choice Behavior: A Theoretical Analysis*. New York, NY: John Wiley & Sons.
- MADDALA, G.S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*. New York, NY: Cambridge University Press.
- MYATT, DAVID P., AND CHRIS WALLACE (1998): "Adaptive Dynamics and Payoff Heterogeneity," laser-script, Nuffield College, Oxford.
- McKELVEY, RICHARD D., AND THOMAS R. PALFREY (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10, 6–38.
- RANKIN, FREDERICK, JOHN VAN HUYCK, AND RAYMOND BATTALIO (2000): "Strategic Similarity And Emergent Conventions: Evidence from Scrambled Payoff Perturbed Stag Hunt Games," *Games and Economic Behavior*, 32, 315–337.
- ROBSON, ARTHUR J., AND FERNANDO VEGA-REDONDO (1996): "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," *Journal of Economic Theory*, 70, 65–92.
- 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.
- SCHMIDT, DAVID, ROBERT SHUPP, JAMES WALKER, AND ELINOR OSTROM (1997): "Playing Safe in Coordination Games: The Role of Risk Dominance, Payoff Dominance, Social History, and Reputation," laser-script.
- STAHL, DALE O. (1996): "Boundedly Rational Rule Learning in a Guessing Game," *Games and Economic Behavior*, 16, 303–330.

- (1999): "Evidence-Based Rules and Learning in Symmetric Normal-Form Games," *International Journal of Game Theory*, 28, 111–130.
- STRAUB, PAUL (1995): "Risk Dominance and Coordination Failure in Static Games," *The Quarterly Review of Economics and Finance*, 35, 339–363.
- VAN HUYCK, JOHN, JOE COOK, AND RAYMOND BATTALIO (1997): "Adaptive Behavior and Coordination Failure," *Journal of Economic Behavior and Organization*, 32, 483–503.
- WEIBULL, JÜRGEN (1995): *Evolutionary Game Theory*. Cambridge, MA: The MIT Press.
- YOUNG, PEYTON (1993): "The Evolution of Conventions," *Econometrica*, 61, 57–84.