Learning Conditional Behavior in Similar Stag Hunt Games

Dale O. Stahl and John Van Huyck

January 2002

Abstract: This paper reports an experiment that varies the range of Stag Hunt games experienced by the participants. The experiment provides evidence that changing the range influences the likelihood of efficient conventions emerging. In the experiment, we observe conditional behavior, much like risk dominance, emerging with experience. We develop a model of conditional expectations to explain these stylized facts that depends crucially on the assumption that after a brief learning period participants categorize their experience using the same relative bandwidth in both treatments even though the range of experience is twice as large in treatment 1 as it is in treatment 2. The assumption can not be rejected by the data. The analysis provides a formal example in which increasing experienced diversity by changing the way similar experiences are categorized increases the likelihood of efficient conventions emerging in communities playing similar Stag Hunt games.

Key Words: Payoff Dominance, Risk Dominance, Similarity, Categorization, Mean Matching, Evolutionary Games.

JEL Classification: c72, c78, c92, d83

Acknowledgments: Nick McKinney and Chris Ball provided able research assistance. Patrick Williams programmed the graphical user interface. The National Science Foundation and Texas Advanced Research Program provided financial support. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the Texas Advanced Research Program.

I. Introduction

Solution concepts that satisfy subgame consistency allow one to construct a general deductive theory of games that is independent of history. The solution of a game is subgame consistent if the solution is independent of the path leading up to the game. Subgame consistent solution concepts have the desirable property that one need not know how a game arose in order to prescribe or predict what will happen.

In general, games have multiple solutions that are neither equivalent nor interchangeable. The class of two player Stag Hunt games, $g(x)$, is an example, see figure 1. Members of the class have three mutually consistent strategy combinations. The strategy combination (β,β) is a payoff dominant equilibrium, while the strategy combination (α,α) is a secure equilibrium. Harsanyi and Selten's (1988) original deductive selection theory selects (β,β), but Harsanyi's (1995) revised theory based on risk dominance selects (β,β) when $x < 0.5$, (α,α) when $x > 0.5$, and the mixed equilibrium when $x = 0.5$.[1]

	α	β
α	x, x	$x, 0$
β	$0, x$	1,1

Figure 1: Two player Stag Hunt Games ($0 < x < 1$)

Without a unique solution it is not obvious even from the point of view of strategic rationality that history shouldn't matter. For example, past experience with the game could condition the players' beliefs in a way that solves the strategy coordination problem arising from multiple strategically rational solutions. Many investigators have found that this strategy coordination problem can be solved by participants if they are given repeated experience with the game, see Battalio, Samuelson, and Van Huyck (2001) for example. Moreover, the equilibrium selected depends on

[1] The equilibrium in mixed strategies is $\{(1-x, x), (1-x, x)\}$. Carlsson and van Damme (1993) also select the risk-dominant equilibrium.

the experience of the community from which the two players are drawn.[2]

Learning models formalize the way past experience with the game influences behavior. Most learning models assume that the same game is played repeatedly. Even if the game has never been encountered in the past, experience with games similar to it could also condition players' beliefs and solve the strategy coordination problem. One of the defining human characteristics is the ability to classify strategic situations and gauge the similarity between situations. Our ability to understand and predict human behavior would be greatly enhanced by a successful theory of how past experiences with similar situations affect current behavior.

Very few theoretical results exist for learning in similar games. LiCalzi (1995) extends the fictitious play dynamic to similar games. He demonstrates that strict equilibria need not be absorbing and that if convergence takes place, what is learned may be path-dependent. LiCalzi's results are for repeated games rather than either the random matching or mean matching protocols discussed below. The myopic best response function assumed in the fictitious play literature leads to wildly inaccurate predictions when applied to repeated games, see almost any paper in the early experimental literature.

Rankin, Van Huyck, and Battalio (RVHB, 2000) report an experiment in which participants play a sequence of similar Stag Hunt games in which payoffs and labels change each period. Each cohort consisted of 8 participants. The participants were randomly pairwise matched and presented with a sequence of 75 Stag Hunt games, $g(x)$. The value of x in period t was independently and uniformly selected from {1/370, 2/370, ..., 369/370}. A constant ε_t was added to all payoffs each period, where ε_t was independently and uniformly selected from {0/370, 1/370, ..., 50/370}. The action labels {A,B} were equally likely to designate {α, β} or {β,α}. These perturbations resulted in payoff dominance emerging as the conventional equilibrium selection principle even when the risk dominant equilibrium had a large basin of attraction under the myopic best response dynamic. There was little evidence for the conditional behavior predicted by risk dominance.[3]

We will refer to the action with the constant $x_t+\varepsilon_t$ payoff as the

---

[2]See also Cooper *et al.* (1992), Straub (1995), Friedman (1996), Clark *et al.* (1996), and Schmidt *et al.* (1997).

[3]The frequency of the payoff dominant action declines significantly for $x_t > 0.8$ rather than $x_t > 0.5$ and continues to account for more than half the choices for values of x_t as large as 0.97.

"Secure" (SEC) action, and we will refer to the other action as the Payoff-Dominant (PD) action.[4] The participants were not told that they would be facing these similar Stag Hunt games: only that they would be facing a sequence of decisions. Thus, the participants were not prompted to look for similarities, and the design masked the similarities by permuting the rows and columns and randomizing the x_t and ε_t values. The RVHB data leave no doubt that the participants perceived the similarity of the games, because they invariably converged to the PD equilibrium, which would have been impossible if they had not identified and labeled the actions as SEC and PD or something equivalent.

This paper reports an experiment that changes the RVHB design in two important ways. The participants are matched against everyone in the cohort each period and receive a payoff equal to the mean of these matches. We will call this protocol mean matching.[5] We report two treatments using mean matching. The first treatment uses the same x_t sequence as RVHB, and the second treatment restricts x_t to be independently and uniformly selected from {185/370, 186/370, ..., 369/370} without replacement. In the second treatment, risk dominance always selects the SEC action.

These changes to the design result in one cohort that converges to the PD convention as in RVHB, two cohorts that converge to the SEC convention, and two cohorts that converge to what we will refer to as conditional behavior. Roughly speaking, in the two cohorts exhibiting conditional behavior the PD action was chosen when $x_t < 0.75$ and the SEC action was chosen when $x_t > 0.75$.

There is no way such behavior could have arisen without the participants recognizing the similarity of the games *and* conditioning on x_t. While the participants may have formed conditional beliefs about what their opponents would do and then chose a best reply to those beliefs, it is also possible that participants merely imitated the population but to an extent that depended on the similarity of the current game and the previous game: that is, on $|x_{t+1} - x_t|$. Such imitative behavior with inertia would generate a *conditional trend* in behavior.

Given these experiment findings, any theory that explains the data will have to incorporate conditional strategies, like the cut point rules in RVHB, conditional trends, and/or conditional beliefs. This paper rejects cut point

---

[4] This action could also be referred to as the Maximax action given $x_t < 1$.

[5] The evolutionary games literature refers to this as 'playing-the-field,' see Crawford (1991) for some important differences between random pairwise matching and playing-the-field.

rules in favor of extending Stahl (1999) to include conditional trends/beliefs to explain the conditional behavior observed in the experiment.

II. Experimental Design

Human participants played a sequence of payoff perturbed 2×2 Stag Hunt games for seventy-five periods with a mean matching protocol. The mean matching protocol gives participants complete information on the current state of their 8 person cohort. Since Random matching only gives feedback on one other participant in the cohort, the two protocols are not directly comparable. Mean matching is closer to the information assumptions made in the learning models we were interested in using to fit the data. Figure 2 is a screen grab illustrating the feedback a subject would receive if they chose action B and the six other participants chose A and one other participant chose action B. Payoffs denote tenths of a cent. The interface recorded their average earnings for the period and their balance in a record grid (not shown).

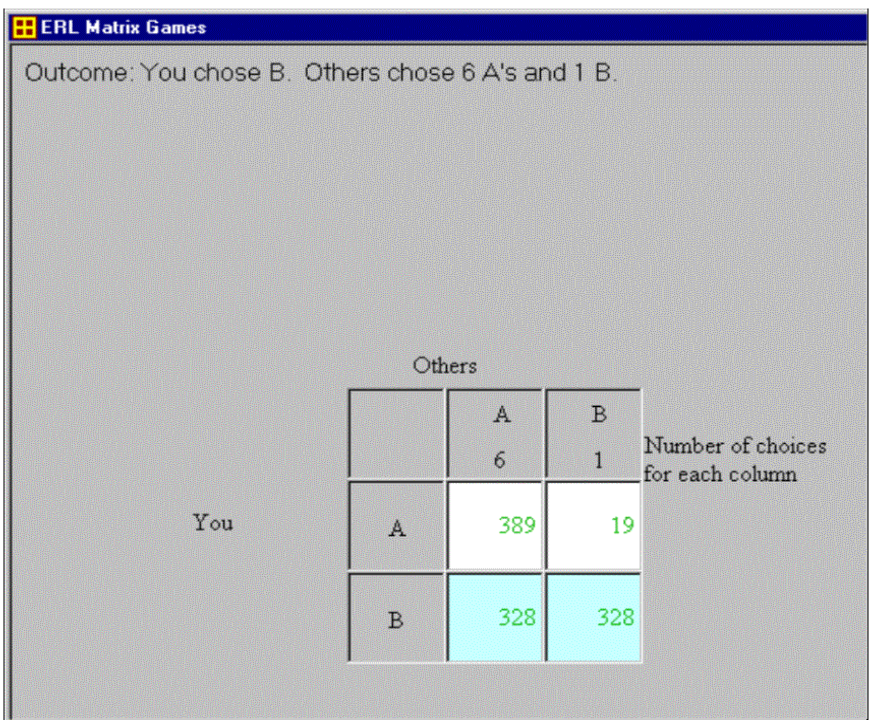

Figure 2: Outcome Screen.

Appendix A gives the seventy-five games used in treatment 1 and 2 respectively. The participants were told the following about the sequence: "Seventy-five earnings tables, one for each of the seventy-five periods, have been generated by a computer. Many sequences of seventy-five earnings tables were generated. One of these sequences will be used in today's session." Thus, the participants were not prompted to look for similarities, and the design masked the similarities by permuting the rows and columns and randomizing the x_t and ε_t values. Figure 3 graphs the x_t values participants actually faced in the two treatments under the own money motivated abstraction assumption.

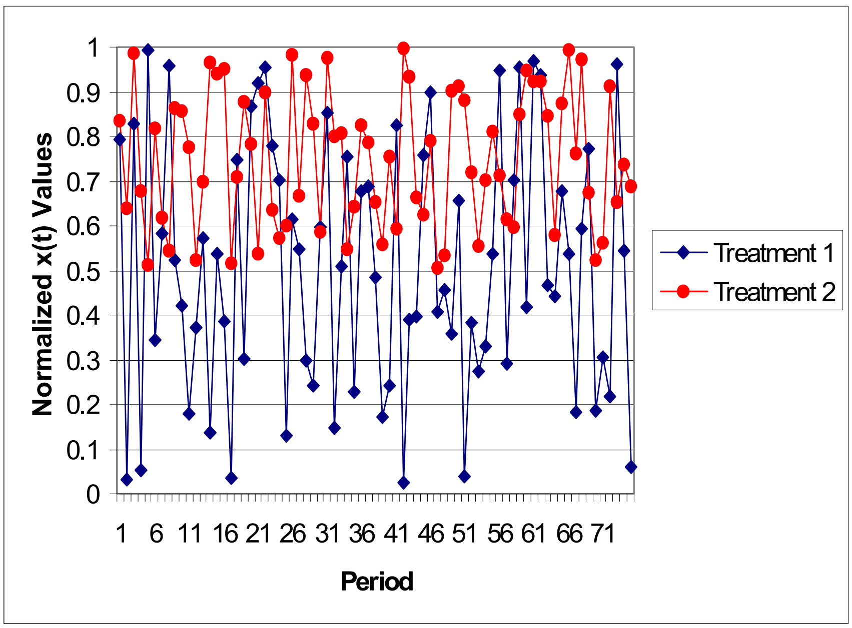

Figure 3 : Normalized $x(t)$ values by treatment.

The participants made choices using a computer assisted graphical user interface available in the TAMU economics research laboratory. No preplay communication was allowed. After each repetition of the period game, participants' earnings were calculated for that period. To further limit the usefulness of retrospective selection principles, there was no computer assisted record keeping of experienced payoff matrices or action combinations. Participants did have access to pencil and paper and could have kept what ever records they felt would be useful. Almost all chose not to keep any records whatsoever.

The participants were undergraduate students at Texas A&M

University. A total of 80 students participated in the ten cohorts reported below. After the instructions, but before the session began, the participants filled out a questionnaire to ensure that it was common information that everyone understood how to compute payoffs for themselves and their opponents. In the seventy-five period sessions, which take about two hours to conduct, a subject could earn as much as $29.50 in treatment 1 and $29.46 in treatment 2.

III. EXPERIMENTAL RESULTS

Initially, there is little difference between the frequency of the PD action in treatment 1 and 2, when risk dominance conflicts with payoff dominance, that is, when x_t is in [0.5,1.0]. In the first 25 periods, 49 percent of the actions are PD in treatment 1 and 45 percent of the actions are PD in treatment 2. This is a difference of only 4 percent. By the last 25 periods, the difference has grown to 30 percent: 79 percent of the actions are PD in treatment 1 and 48 percent of the actions are PD in treatment 2. Having experience in games in which risk dominance reinforced payoff dominance, treatment 1, influenced behavior in those games in which payoff dominance conflicted with risk dominance.

A useful way to display the data is with histograms of the frequency of the PD action for the first and last 25 periods. Table 1 reports the fixed bin histograms of the PD action frequency by cohort. Cohorts 1 to 5 used treatment 1 and include observations on x_t in [0.0,0.5]. Cohorts 6 to 10 used treatment 2 and do not include observations on x_t in [0.0,0.5].

Playing the PD action was always a best response to the population state when x_t was in [0.0,0.5] both in the first twenty-five periods and the last twenty-five periods of treatment 1. By the end of the session cohorts 2, 3, and 4 implement the PD equilibrium in every game in which risk-dominance selects the PD equilibrium, that is, x_t in [0.0,0.5]. Cohorts 1 and 5 miss the PD equilibrium by a tremble, see Table 1.

<Table 1 about here.>

Things are more complicated when the participants encountered games in which risk dominance conflicts with payoff dominance. For the bin with x_t in [0.875,1.0], playing the PD action was never a best response to the population state initially. During the last 25 periods the PD action was a best response in cohorts 1, 3, and 4 of treatment 1 and cohort 6 of treatment 2 when x_t was in [0.0,0.875]. The fact that the PD action was a best response to the population state at the end of the session three times as often in treatment 1 as in treatment 2 suggests that restricting the range of x experienced by the cohort such that risk dominance always conflicts with payoff dominance lowers the likelihood of the cohort converging to the PD equilibrium. Of the twenty cells per treatment with $x_t > 0.5$ observe that 17 of 20 cells in treatment 1 have more than 75 percent PD choices, while only 8 of 20 cells in treatment 2 have more than 75 percent PD choices and 10 of 20 cells have less than 25 percent PD choices.

While the behavior of cohorts 7 and 8 is reminiscent of the results from sessions using exactly the same game repeatedly namely convergence to the inefficient secure equilibrium, cohorts 9 and 10 clearly exhibit conditional behavior. The rule play PD when $x_t < 0.75$, and to play SEC otherwise is a best response to the states experienced by cohorts 9 and 10 by the end of the session. This is indisputable evidence of behavior that depends on x_t.

RVHB introduced cutpoint rules as a way to fit participants' behavior. The participant compares x_t to a cutpoint level, c, and tends to choose the PD action when $x_t < c$. We estimated a rule learning model as in Stahl (2000) with a space of cut point rules. However, the maximum likelihood estimate of the mean cutpoint was always 1, which we don't feel captures the conditional behavior observed in the data: especially, the way this conditional behavior was influenced by the treatments. Thinking about cut point rules lead us to a different approach. A cutpoint rule is a best reply to a belief that others will use that cutpoint rule. Given that the participants appear to recognize the strategic similarity of the games, which differ only by the value of x_t , it is not unreasonable to suppose that the participants would and should form conditional beliefs. We take up this approach next.

IV. DATA ANALYSIS

In order to form conditional beliefs participants must categorize the games into conditioning states. Categorization provides the gateway between perception and cognition. It is because treatment 1 and treatment 2 might result in different categorizations of experience that we expect the design to influence the emergent convention. The categorization ability of humans have usually been found to exceed the categorization ability of algorithms, see Barsalou (1992). So, the task of accurately modeling conditional beliefs in this environment is non-trivial.

If there were a small number of conditioning states, say 2 to 5, and the states were independent, then one could compile various statistics for each state by itself. However, the environment we are studying entails 369 states for x_t in treatment 1 and 184 states for x_t in treatment 2. If subjects don't categorize nearby states as similar, then they could not coordinate on a convention as effectively as they do. Indeed, since only 75 values of x_t are observed by a participant throughout the whole experiment, the assumption of correlated behavior for nearby states is crucial to being able to form conditional beliefs for all the states. Moreover, at the beginning of the experiment, the participants do not know the range of x_t, but must learn that from experience as well.

In the absence of a statistically optimal solution, we draw on tradition, statistics, psychology and neuroscience. From tradition, we adopt the standard partial adjustment dynamic which has worked well in environments with identical games, and we extend this to environments with similar games. From psychology and neuroscience, we hypothesize that humans partition the perceived range of x_t into a small number of bins into which experience is gathered. From neuroscience and statistics, we hypothesize a bandwidth smoothing adaptive process.

A. Conditional Partial Adjustment Model

Both trend formation via imitation and inertia and simple adaptive belief formation can be modeled with a generic partial adjustment dynamic. With just two actions (SEC and PD), it suffices to express trends and beliefs as the proportion of participants choosing the PD action, which we will denote as $q^i(x_t)$. Letting $p^{-i}(x_t)$ denote the actual proportion of *other* participants choosing the PD action in period t, the partial adjustment dynamic is given by

$$q^i\,(x_t) \;=\; \theta_t\, p^{-i}(x_{t-1}) + (1\text{-}\theta_t)\, q^i(x_{t-1}) \qquad t > 1, \qquad (1)$$

with $q^i(x_1) = 0.5$. As a model of trend formation, θ_t represents the probability of imitation, with $(1\text{-}\theta_t)$ being the probability of sticking with last period's trend. As a model of adaptive expectations, θ_t represents the weight on current evidence and $(1\text{-}\theta_t)$ is the weight on prior expectations.

The subscript on θ_t serves to indicate potential dependence on distance $|x_t - x_{t\text{-}1}|$. For example, consider

$$\theta_t \;=\; \theta\; exp\{-0.5[(x_t - x_{t\text{-}1})/s]^2\}. \qquad (2)$$

If all games were identical ($x_t - x_{t-1} = 0$), then θ_t would be a constant θ. For non-identical games, $\theta_t < \theta$, implying that the probability of imitation decreases with distance, and/or recent observations have less effect on the current belief due to the difference between the situations.

The parameter s is a scaling parameter that gives behavioral meaning to the distance $| x_t - x_{t-1} |$. We will call s the *bandwidth* of the adaptive expectation process. Distances of more than one bandwidth have significantly less influence. If one were to partition the range of x_t into bins of equal influence, then the bin widths would be proportional to s. In trading off the benefits of finer partitions and accuracy of a histogram, the common practice is to let the number of bins grow as the square root of the aggregate sample size, which translates to letting the bandwidth s decline as the square root of the number of periods. However, practical considerations usually put an upper bound on the number of bins: hence, a lower bound on the bandwidth. It is natural to measure distance and bandwidth relative to the range of x_t.

For these reasons, we hypothesize a time-varying relative bandwidth of the form:

$$s(t) = R \max\{r\, \sigma_1\, t^{-1/2}, \sigma_0\}. \tag{3}$$

where R is the range of x_t, and r is the ratio of the initial perceived range to the actual range. Since the minimum bandwidth σ_0 does not bind until t becomes sufficiently large, and since the participants in the experiment can easily learn the range from experience, eventually the ratio of the perceived range to the actual range is one, so there is no need for a multiplier in front of σ_0. At the beginning of the experiment, however, the participants do not know the range. Nonetheless, [0, 1] is eminently reasonable as the initial perceived range since the minimum and maximum payoffs of the first (and every) game are 0 and 1 respectively. This perceived range is strongly reinforced in Treatment 1 in which the first five draws of x_t were {0.792, 0.032, 0.830, 0.054, 0.995}; and of course there is never any contrary evidence. Therefore, for Treatment 1, we assume $r = 1$.[6]

For Treatment 2, we again argue that [0, 1] is eminently reasonable as the initial perceived range, but since the first five draws of x_t were {0.835, 0.638, 0.986, 0.678, 0.511}, most participants will have come to believe that $x_t < 0.5$ is unlikely. By setting $r = 2$, equation (3) captures this effect with one set of (σ_0, σ_1) coefficients for both treatments, that is, for Treatment 2, the initial bandwidth is $2\sigma_1$ relative to the true range [0.5, 1], and by the square root relationship, after four periods, the bandwidth

6 This assumption is really an identification restriction since only the product $r\, \sigma_1$ can be identified.

declines to σ_1 relative to the true range (as if the true range becomes known).

As a benchmark, we estimated this conditional partial adjustment model on the individual participant data. When analyzing this data we will always fit models by maximizing the likelihood of the *individual* choices rather than the aggregated choices. Specifically, we compute the likelihood of each participant's 75 choices, and then aggregate over participants. The MLE results are given in table 2:

Table 2: Maximum likelihood estimates of CAE model

	θ	σ_0	σ_1	LL
Pooled	0.532	0.128	0.498	-2189.13
Treatment 1	0.530	0.127	0.581	-985.81
Treatment 2	0.514	0.117	0.464	-1201.98

Summing the log-likelihood (LL) values of each treatment and subtracting the pooled log-likelihood value gives 1.34; twice this is distributed chi-squared with six degrees of freedom, and has a p-value of 0.969. Thus, we cannot reject the pooled parameter estimates. Moreover, since $r = 2$ was imposed on the pooled estimates, we cannot reject that specification.

Note that if distance were measured on a scale independent of the range of x_t, then we should have found the estimate of σ_0 for Treatment 1 to be twice the estimate for Treatment 2, and we should have rejected the pooled estimate which measures distance relative to the range of x_t. Therefore, a major result is that we cannot reject the hypothesis that *distance is measured relative to the objective range* of x_t with the adjustment for initial perceptions as specified.

The units for (σ_0, σ_1) are expressed as a proportion of the range. Thus, the initial bandwidth is about half the range[7], while the asymptotic minimum bandwidth is about one-eighth of the range. The latter is essentially equivalent to dividing the range into eight bins, which is consistent with psychometric evidence that humans can accurately discriminate between approximately seven categories.

B. Logit Best Reply with Inertia and Conditional Adaptive Expectations

7 The whole range for Treatment 2, since the range is initially misperceived to be [0, 1].

The logit best-reply with inertia and adaptive expectations (LBRIAE) model of Stahl (1999) can be adapted to incorporate conditional expectations. That model supposes that with probability δ the player follows the "herd", and with probability (1-δ) the player makes a logit best reply to an adaptive expectation of the other player's choice. Following the herd means behaving as a first-order trend process. In the present context, this means that behavior is given by the conditional trend defined by eqs(1-3). Then, the probability of choosing the PD action is

$$\delta\, q^i(t|x_t) + (1\text{-}\delta)\ \mathrm{br}[q^i(t|x_t), \nu],$$

where $\mathrm{br}[q^i(t|x_t), \nu]$ is the logit best-reply function with precision ν. To this, we add a tendency to tremble η, so the ultimate probability of choosing the PD action conditional on x_t is

$$P^i(x_{t+1};\nu, \delta, \eta) \equiv (1\text{-}\eta)\{\delta\, q^i(t|x_{t+1}) + (1\text{-}\delta)\ \mathrm{br}[q^i(t|x_{t+1}), \nu]\} + \eta/2. \qquad (4)$$

We call this the Logit Best-Reply with Inertia and Conditional Adaptive Expectations (LBRICAE) model. The MLE results are given in table 3:

Table 3: Maximum Likelihood Estimates of LBRICAE model

Data	ν	δ	θ	σ_0	σ_1	η	**LL**
Pooled	12.9	0.861	0.663	0.142	0.524	0.012	-2147.26
Treatment 1	9.41	0.689	0.509	0.182	0.417	0.003	-951.18
Treatment 2	47.2	0.920	0.737	0.111	0.479	0.020	-1181.86

The estimates of δ are quite high, particularly for Treatment 2, indicating that the conditional trend component accounts for most of the data. Nonetheless, the improvement in the LL over simply conditional adaptive expectations is very statistically significant. For instance, the increase for the pooled data is 41.87, which doubled is distributed chi-squared with 6 degrees of freedom and has a *p*-value less than 4.82E-18. Thus, we can strongly reject the conditional partial adjustment process alone in favor of the LBRICAE model. The parameters of the conditional partial adjustment process (θ, σ_0, σ_1) have changed substantially both across treatments and from the addition of logit best-replies. Furthermore, we now reject the pooled estimates (p-value = 0.005), so this is not a model that can explain the data for both treatments with one set of parameters.

C. Worldly Type with Inertia and Conditional Adaptive Expectations

One of the limitations of the LBRICAE model is the implicit assumption that the logistic best responders believe everyone else will follow the herd. In contrast, a logistic best responder could also believe that some other players will choose the best-reply to $q^i(t|x_{t+1})$, while others will deviate to the PD (or Maximax) action, leading to a compound belief:

$$Q^i(t;\, \alpha_1, \alpha_2|x_{t+1}) \equiv \alpha_1\, q^i(t|x_{t+1}) + \alpha_2\, \mathrm{br}[q^i(t|x_{t+1}),\infty] + (1\text{-}\alpha_1\, \text{-}\alpha_2)\,. \qquad (5)$$

We then assume a logit best reply to this belief: $\mathrm{br}[Q^i(t;\, \alpha_1, \alpha_2|x_{t+1}),\, \nu]$ with precision ν. We define $\nu_1 \equiv \alpha_1\, \nu$, $\nu_2 \equiv \alpha_2\, \nu$, and $\nu_3 \equiv (1\text{-}\alpha_1\, \text{-}\alpha_2)\nu$, and estimate $(\nu_1, \nu_2, \nu_3, \delta, \theta, \sigma_0, \sigma_1, \eta)$. We will designate this model as WICAE (Worldly with Inertia and Conditional Adaptive Expectations). The MLE results are given in table 4 with the exception of ν_2 whose MLE was exactly zero:

Table 4: Maximum Likelihood Estimates of WICAE model.

Data	ν_1	ν_3	δ	θ	σ_0	σ_1	η	LL
Pooled	8.20	7.83	0.726	0.601	0.157	0.621	0.008	-2107.31
Treatment 1	6.84	3.29	0.611	0.469	0.185	0.899	0.003	-939.62
Treatment 2	15.1	17.9	0.790	0.694	0.152	1.17	0.010	-1157.39

The first result is that the LL is dramatically increased over the LBRICAE model; hence, the inclusion of PD (or Maximax) beliefs helps explain the data.[8] Solving for the weight on PD beliefs in equation (5) gives a weight of 0.48, which is large. The second result is that level-2 beliefs are given zero weight. The third result is that we cannot reject the pooled estimates (p-value = 0.116). Thus, this model is a candidate for explaining the observed behavior in both treatments with one set of parameters.

D. Goodness-of-Prediction

As another measure of how well this model can predict, we use the

---

8 A related model in which behavior is a mixture of LBRICAE and a PD type was considered, but it does not fit the data nearly as well as the WICAE (pooled LL = -2135.50) and contains one more parameter.

MLE parameter estimates to simulate 75-period paths for 1000 eight-participant cohorts. For each simulated path, we compute the smoothed proportion of PD play for the last 25 periods just as we did for the actual data. From the actual data, three categories are interesting: (1) those paths for which the smoothed proportion of PD play for the last 25 periods exceeds 0.75; (2) those paths for which the difference between the low x bin and high x bin exceeds 0.50 (indicating a substantial negative slope to the conditional trend/belief); and (3) those paths for which the smoothed proportion of SEC play exceeds 0.75. Table 5 reports the percentage of paths with these characteristics for the actual data and simulated data by treatment. Note that not all paths fall into these three categories, so the percentages do not add to 100.

Table 5: Model simulations versus observed behavior.

	Treatment 1			**Treatment 2**		
	PD > 0.75	Diff. > 0.5	SEC > 0.75	PD > 0.75	Diff. > 0.5	SEC > 0.75
Actual Data	40.0%	20.0%	0.0%	20.0%	40.0%	20.0%
CAE - pooled	7.1%	0.9%	6.5%	16.4%	2.3%	12.4%
LBRICAE - pooled	2.6%	21.1%	10.7%	0.1%	2.6%	85.6%
WICAE - pooled	22.1%	10.7%	0.0%	0.2%	77.6%	2.5%

We can draw the following inferences from this table. First, without the tilt towards the PD action embodied in the WICAE model, SEC choices are over-predicted especially for Treatment 2. Second, the WICAE model predicts that substantial conditional behavior (Difference > 0.5) will be twice as likely (or more) in Treatment 2 than in Treatment 1, as found in the actual data. Third, for Treatment 1 only the WICAE model predicts more instances of essentially unconditional PD choices than either unconditional SEC choices or substantial conditional behavior.

V. Discussion and Conclusions

Experience with a greater range of Stag Hunt games increases the likelihood that a convention based on payoff dominance emerges in a laboratory community. Specifically, when x_t ranged between 0 and 1 a convention based payoff dominance emerged three times as often as when x_t ranged between 0.5 and 1. The likelihood of emergent conditional behavior was also influenced by the range of Stag Hunt games experienced.

This paper develops a model of conditional adaptive expectations to fit the laboratory phenomena. The key analytical assumption is that after a brief learning period participants categorize their experience using the same bandwidth in both treatments even though the range of experience is twice as large in treatment 1 as in treatment 2. An important result is that this assumption can not be rejected by the data.

Simply including conditional adaptive expectations into a logistic best reply model did not allow us to explain the data with one set of fitted parameters. It was necessary to introduce an exogenous belief in the salience of the PD action.[9] Again we could not reject the key analytical assumption that after a brief learning period participants categorize their experience using the same relative bandwidth in both treatments even though the range of experience is twice as large in treatment 1 as in treatment 2. Moreover, simulations with the fitted parameters reproduce the stylized facts described above. The estimated minimum bandwidth is about one-seventh the range of x_t, which is consistent with the stylized categorization facts reported by cognitive psychologists. We have thus developed a formal example in which increasing experienced diversity increases the likelihood of efficient conventions emerging in laboratory communities playing similar Stag Hunt games.

---

[9] This bias towards efficiency seems similar to the bias remarked on by Van Huyck, Cook, and Battalio (1997) and Van Huyck, Battalio, and Rankin (2001). Haruvy and Stahl (2000) also find a bias towards efficiency that can be attributed to maximax behavior rather than payoff dominance. Van Huyck, Battalio, and Beil (1990) provide an early example in which participants do not report mutually consistent beliefs and, hence, can not be imposing the mutual consistency condition implicit in the payoff dominance equilibrium selection principle.

REFERENCES

Barsalou, L.W., *Cognitive Psychology: An overview for cognitive scientists*, (Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1992)

Battalio, R., L. Samuelson, and J. Van Huyck, (2001) "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games," *Econometrica*, 69(3), 749-764.

Carlsson, H. and Eric van Damme, (1993) "Global games and equilibrium selection," *Econometrica* 61, 989-1018.

Clark, K., Stephen Kay and Martin Sefton, (1996), "When are Nash Equilibria Self-Enforcing? An Experimental Analysis."

Cooper, R., D.V. DeJong, R. Forsythe, and T.W. Ross, (1992) "Communication in coordination games," *Quarterly Journal of Economics* 107, 739-773.

Crawford, V., (1991), "An 'Evolutionary' Interpretation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination," *Games and Economic Behavior* 3, 25-59

Gilboa, I. and D. Schmeidler, (1995) "Case-based Decision Theory," *Quarterly Journal of Economics*, 110, 605-639.

Friedman, D., (1996) "Equilibrium in evolutionary games: Some experimental results," *Economic Journal* 106, 1-25.

Fudenberg, D. and Kreps, D. (1993) "Learning Mixed Equilibria," *Games and Economic Behavior*, **5**, 320-367.

Harsanyi, J.C. (1995) "A new theory of equilibrium selection for games with complete information," *Games and Economic Behavior* 8, 91-122.

Harsanyi, J.C. and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*, Cambridge,MA: MIT Press.

Haruvy, E. and D.O. Stahl, (2000) "Initial Play and Equilibrium Selection in Symmetric Normal-Form Games."

LiCalzi, M. (1995). "Fictitious Play by Cases," *Games and Economic Behavior*, **11**, 64-89.

Rankin, F., J.B. Van Huyck, and R.C. Battalio, (2000) "Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games," *Games and Economic Behavior*, 32, 315-337.

Schmidt, D., Robert Shupp, James Walker, Elinor Ostrom, (1997) "Playing Safe in Coordination Games: The Role of Risk Dominance, Payoff Dominance, Social History, and Reputation."

Stahl, D.O. (1999), "A Horse Race Among Action Reinforcement Learning Models."

Stahl, D.O. (2000), "Rule Learning in Symmetric Normal-Form Games: Theory and Evidence," *Games and Economic Behavior,* 32(1), 105-138.

Straub, P., (1995) "Risk Dominance and Coordination Failure in Static Games," *The Quarterly Review of Economics and Finance* 35, 339-363.

Van Huyck, J.B., J.P. Cook, R.C. Battalio, (1997) "Adaptive Behavior and Coordination Failure," *Journal of Economic Behavior and Organization*, 32, 483-503.

Van Huyck, J.B., R.C. Battalio, and R.O. Beil, (1990) "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80(1), 234-248.

Van Huyck, J.B., R.C. Battalio, and F.W. Rankin, (2001) "Evidence on Learning in Coordination Games ."

Appendix

Treatment 1: $g(x_t)$ for x_t in [0,1]

```
M01 x01=0.79
 299,299
   6,376
M02 x02=0.03
 415, 45
  57, 57
M03 x03=0.83
 313,313
   6,376
M04 x04=0.05
 407, 37
  57, 57
M05 x05=0.99
 415,415
  47,417
M06 x06=0.34
 149,149
  22,392
M07 x07=0.58
 256,256
  41,411
M08 x08=0.96
 356,356
   2,372
M09 x09=0.52
 242,242
  48,418
M10 x10=0.42
 389, 19
 175,175
M11 x11=0.18
 108,108
  42,412
M12 x12=0.37
 371,  1
 138,138
M13 x13=0.57
 229,229
  18,388
M14 x14=0.14
 404, 34
  84, 84
M15 x15=0.54
 397, 27
 225,225
M16 x16=0.39
 400, 30
 173,173
M17 x17=0.04
  62, 62
  49,419
M18 x18=0.75
 315,315
  38,408
M19 x19=0.30
 370,  0
 112,112
M20 x20=0.87
 330,330
   9,379
M21 x21=0.92
 377,377
  37,407
M22 x22=0.95
 388,388
  35,405
M23 x23=0.78
 320,320
  32,402
M24 x24=0.70
 419, 49
 309,309
M25 x25=0.13
  77, 77
  29,399
M26 x26=0.61
 390, 20
 247,247
M27 x27=0.55
 234,234
  31,401
M28 x28=0.30
 373,  3
 114,114
M29 x29=0.24
 398, 28
 117,117
M30 x30=0.60
 384, 14
 235,235
M31 x31=0.85
 322,322
   6,376
M32 x32=0.15
  85, 85
  31,401
M33 x33=0.51
 237,237
  49,419
M34 x34=0.75
 400, 30
 309,309
M35 x35=0.23
 419, 49
 134,134
M36 x36=0.68
 269,269
  18,388
M37 x37=0.69
 416, 46
 300,300
M38 x38=0.48
 186,186
   7,377
M39 x39=0.17
  96, 96
  32,402
M40 x40=0.24
 120,120
  31,401
M41 x41=0.82
 370,  0
 305,305
M42 x42=0.02
  38, 38
  29,399
M43 x43=0.39
 416, 46
 190,190
M44 x44=0.40
 390, 20
 167,167
M45 x45=0.76
 377,  7
 287,287
M46 x46=0.90
 348,348
  16,386
M47 x47=0.41
 384, 14
 165,165
M48 x48=0.46
 195,195
  26,396
M49 x49=0.36
 394, 24
 156,156
M50 x50=0.66
 252,252
   9,379
M51 x51=0.04
  17, 17
   3,373
M52 x52=0.38
 150,150
   9,379
M53 x53=0.27
 414, 44
 145,145
M54 x54=0.33
 145,145
  23,393
M55 x55=0.54
 199,199
   0,370
M56 x56=0.95
 392, 22
 372,372
M57 x57=0.29
 419, 49
 157,157
M58 x58=0.70
 417, 47
 307,307
M59 x59=0.95
 391, 21
 374,374
M60 x60=0.42
 391, 21
 175,175
M61 x61=0.97
 400,400
  42,412
M62 x62=0.94
 366,366
  19,389
M63 x63=0.47
 404, 34
 207,207
M64 x64=0.44
 171,171
   8,378
M65 x65=0.68
 283,283
  32,402
M66 x66=0.54
 372,  2
 201,201
M67 x67=0.18
 412, 42
 109,109
M68 x68=0.59
 259,259
  39,409
M69 x69=0.77
 297,297
  12,382
M70 x70=0.19
 404, 34
 103,103
M71 x71=0.31
 150,150
  37,407
M72 x72=0.22
 395, 25
 106,106
M73 x73=0.96
 367,367
  11,381
M74 x74=0.54
 400, 30
 231,231
M75 x75=0.06
 415, 45
  67, 67
```

Treatment 2: $g(x_t)$ for x_t in [0.5,1.0]

```
M01 x01=0.84
 389, 19
 328,328
M02 x02=0.64
 253,253
  17,387
M03 x03=0.99
 397,397
  32,402
M04 x04=0.68
 277,277
  26,396
M05 x05=0.51
 405, 35
 224,224
M06 x06=0.82
 303,303
   1,371
M07 x07=0.62
 257,257
  28,398
M08 x08=0.54
 221,221
  20,390
M09 x09=0.86
 377,  7
 326,326
M10 x10=0.86
 360,360
  43,413
M11 x11=0.78
 394, 24
 311,311
M12 x12=0.52
 217,217
  24,394
M13 x13=0.70
 404, 34
 293,293
M14 x14=0.96
 374,  4
 361,361
M15 x15=0.94
 357,357
   9,379
M16 x16=0.95
 382, 12
 364,364
M17 x17=0.52
 220,220
  29,399
M18 x18=0.71
 406, 36
 298,298
M19 x19=0.88
 402, 32
 356,356
M20 x20=0.78
 310,310
  21,391
M21 x21=0.54
 388, 18
 216,216
M22 x22=0.90
 341,341
   8,378
M23 x23=0.64
 415, 45
 280,280
M24 x24=0.57
 417, 47
 259,259
M25 x25=0.60
 224,224
   2,372
M26 x26=0.98
 413,413
  50,420
M27 x27=0.67
 410, 40
 287,287
M28 x28=0.94
 393,393
  47,417
M29 x29=0.83
 397, 27
 333,333
M30 x30=0.59
 388, 18
 235,235
M31 x31=0.98
 371,  1
 362,362
M32 x32=0.80
 307,307
  11,381
M33 x33=0.81
 388, 18
 317,317
M34 x34=0.55
 411, 41
 243,243
M35 x35=0.23
 268,268
  30,400
M36 x36=0.82
 338,338
  33,403
M37 x37=0.79
 327,327
  36,406
M38 x38=0.65
 272,272
  30,400
M39 x39=0.56
 245,245
  38,408
M40 x40=0.75
 378,  8
 287,287
M41 x41=0.59
 228,228
   9,379
M42 x42=1.00
 382, 12
 381,381
M43 x43=0.93
 375,  5
 350,350
M44 x44=0.66
 405, 35
 280,280
M45 x45=0.62
 381, 11
 242,242
M46 x46=0.79
 319,319
  27,397
M47 x47=0.51
 405, 35
 222,222
M48 x48=0.53
 394, 24
 221,221
M49 x49=0.90
 382, 12
 346,346
M50 x50=0.91
 345,345
   7,377
M51 x51=0.88
 381, 11
 337,337
M52 x52=0.72
 384, 14
 280,280
M53 x53=0.55
 220,220
  15,385
M54 x54=0.70
 394, 24
 284,284
M55 x55=0.81
 334,334
  34,404
M56 x56=0.71
 301,301
  37,407
M57 x57=0.61
 375,  5
 232,232
M58 x58=0.60
 406, 36
 257,257
M59 x59=0.85
 356,356
  42,412
M60 x60=0.95
 380, 10
 361,361
M61 x61=0.92
 373,  3
 344,344
M62 x62=0.92
 409, 39
 381,381
M63 x63=0.85
 385, 15
 328,328
M64 x64=0.58
 378,  8
 222,222
M65 x65=0.87
 335,335
  12,382
M66 x66=0.99
 407, 37
 404,404
M67 x67=0.76
 304,304
  22,392
M68 x68=0.97
 382,382
  23,393
M69 x69=0.67
 282,282
  33,403
M70 x70=0.52
 400, 30
 224,224
M71 x71=0.56
 213,213
   5,375
M72 x72=0.91
 351,351
  14,384
M73 x73=0.65
 372,  2
 243,243
M74 x74=0.74
 303,303
  31,401
M75 x75=0.69
 282,282
  27,397
```

Table 1: Fixed Bin Histograms of PD action Frequency.

	[0.0, 0.125]	(0.125, 0.25]	(0.25, 0.375]	(0.375,0.5]	(0.5,0.625]	(0.625,0.75]	(0.75,0.875]	(0.875,1.0]
				Period 1 to 25				
Cohort 1	0.79	0.88	0.96	1.00	0.91	0.81	0.75	0.72
Cohort 2	0.79	0.88	0.75	0.69	0.38	0.25	0.22	0.09
Cohort 3	0.88	1.00	0.96	0.94	0.91	0.94	0.56	0.69
Cohort 4	0.79	0.79	0.46	0.56	0.22	0.00	0.13	0.03
Cohort 5	0.83	0.92	0.75	0.88	0.63	0.75	0.38	0.53
Cohort 6	.	.	.	.	0.94	0.80	0.79	0.79
Cohort 7	.	.	.	.	0.39	0.45	0.27	0.17
Cohort 8	.	.	.	.	0.34	0.35	0.31	0.21
Cohort 9	.	.	.	.	0.45	0.45	0.13	0.04
Cohort 10	.	.	.	.	0.69	0.65	0.48	0.17
				Period 51 to 75				
Cohort 1	1.00	1.00	1.00	0.97	1.00	1.00	1.00	0.98
Cohort 2	1.00	1.00	1.00	1.00	0.94	0.75	0.75	0.13
Cohort 3	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.78
Cohort 4	1.00	1.00	1.00	1.00	0.97	0.81	1.00	0.43
Cohort 5	1.00	0.96	1.00	1.00	0.94	0.88	0.75	0.68
Cohort 6	.	.	.	.	1.00	1.00	1.00	0.96
Cohort 7	.	.	.	.	0.44	0.20	0.18	0.07
Cohort 8	.	.	.	.	0.15	0.07	0.00	0.04
Cohort 9	.	.	.	.	0.98	0.93	0.23	0.05
Cohort 10	.	.	.	.	0.98	0.95	0.33	0.14