Risk Dominance, Payoff Dominance and Probabilistic Choice Learning

Raymond Battalio, Larry Samuelson, and John Van Huyck

November 1997
Draft c4_1_5

Comments to john.vanhuyck@tamu.edu
Related research available at
http://econlab10.tamu.edu/JVH_gtee/

Abstract: This paper reports an experiment comparing three stag hunt games that have the same best-response correspondence. The games have the same expected payoff from the mixed equilibrium, but differ in the pecuniary incentive a player has to play a best response to other mixtures. In each game, risk dominance conflicts with payoff dominance and selects an inefficient pure strategy equilibrium. We find statistically and economically significant evidence that the expected earnings difference helps explain observed behavior.

Key Words: Payoff dominance, risk dominance, probabilistic choice, exponential fictitious play, bounded rationality, random utility, logistic response equilibria, human behavior.

JEL Classification: c72, c78, c92, d83.

Acknowledgements: We thank Bill Rankin and Nick Rupp for research assistance, Simon Anderson and Richard McKelvey for helpful discussions, Dan Friedman, Robert Forsythe, Paul Straub, Martin Sefton, and their collaborators for making their data available to us. Eric Battalio implemented the experimental design on the TAMU economics laboratory network. The National Science Foundation and the Texas Advanced Research Program provided financial support.

1 INTRODUCTION

The abstraction assumptions specifying feasible strategies, preferences over consequences, and substantively rational players, which define a strategic form game, construct a powerful analytical framework within which to analyze strategic behavior. These abstraction assumptions in turn can be summarized by the best-response correspondence. One need only know the best-response correspondence of a strategic form game to identify its Nash equilibria. The vast majority of the equilibrium concepts in the refinements literature similarly depend only on the best-response correspondence, as do many of the concepts that choose between strict Nash equilibria. All of the information required when applying such an equilibrium concept to a game is then contained in the structure of the game's best responses.

This paper reports an experimental investigation of three games that have the same best-response correspondence, as well as similar payoff magnitudes, but which produce different behavior. Games *2R*, *R*, and *0.6R*, shown in Figures 1, 2, and 3, were used in the experiment. Cell entries denote cents. Under the assumption that players maximize their earnings, each game has two pure-strategy equilibria, including the payoff-dominant equilibrium *(X,X)* and the risk-dominant equilibrium *(Y,Y)*.[1] Each game also has a mixed equilibrium in which *X* is played with probability k^*, where k^* equals 0.8. Games with this structure are commonly referred to as (two player) *stag hunt games*.

The best response correspondence in games *2R*, *R* and *0.6R* is completely determined by k^*. Strategy *X* is a strict best response to any mixture that attaches a probability greater than k^* to *X*, while *Y* is a strict best response to any mixture attaching a lower probability to *X*. To the extent possible, the games also involve payoffs of similar magnitudes. In particular, the expected payoff from the mixed equilibrium is 36 for all three games.

Classical theories, based on models of substantively rational agents, typically treat games *2R*, *R* and *0.6R* identically. 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. Anderlini (1990) and Harsanyi and Selten (1988) choose the payoff-dominant equilibrium. In each case, the prediction does not depend upon

[1] The concepts of payoff dominance and risk dominance are taken from Harsanyi and Selten (1988). The risk-dominance equilibrium in a 2×2 symmetric game is the one with the larger basin of attraction under best-response dynamics.

which of game *2R*, *R* or *0.6R* is played.[2]

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

Figure 1: *2R* Game Form

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

Figure 2: *R* Game Form

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

Figure 3: *0.6R* Game Form

The recent shift from models based on substantive rationality to models of boundedly rational agents has directed attention to learning-based theories of equilibrium selection. Theories based on deterministic dynamics, such as the replicator dynamic or myopic best response dynamic, choose either the payoff-dominant or risk-dominant equilibrium, depending upon which basin of attraction contains the initial condition, where the *separatrix* divides the state space into basins of attraction.[3] Such theories predict history-dependent equilibrium selection. The three games have the same separatrix, k^*, under either the replicator or myopic best response dynamic. If the initial state is on the same side of k^*, the dynamic picks the same equilibrium in all three games.[4]

---

[2] There are other theories that do not choose between the strict Nash equilibria of the stag hunt game, but which do not distinguish between games with the same best-response correspondence. Hillas (1990) introduces an equilibrium concept that relies on the best-response correspondence in a particularly obvious way.

[3] Weibull (1995) contains a good introduction to the replicator dynamic.

[4] Van Huyck, Cook, and Battalio (forthcoming) use an analogy between the separatrix and the continental divide to explain the barrier dividing adaptive behavior in their experiment: "Mark Twain (1962, p.86) describes a remarkable spring at the summit of a Rocky Mountain pass that 'spent its water through two outlets and set it in opposite directions.' One of the streams starts a journey westward to the Gulf of California and the Pacific Ocean. The other starts a journey eastward to the Gulf of Mexico and the Atlantic Ocean. Our search was for a spring that straddles a barrier dividing a continent of human behavior." In the experiment reported in this paper, we were searching for separatrix crossings and so we choose an extreme value of k^* in the hope that all initial conditions will be in the risk dominant basin of attraction.

Kandori, Mailath and Rob (1993) and Young (1993) present stochastic models in which the risk-dominant equilibrium is always selected. Robson and Vega-Redondo (1996) examine a similar model in which the payoff-dominant equilibrium is selected. Despite their different outcomes, each of these theories selects the same equilibrium for all three games.

Our analysis of games *2R*, *R* and *0.6R* is motivated by the observation that the pecuniary incentive to select a best response is always 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 refer to this incentive, given by the difference in the payoffs of the best response and the inferior response, as the *earnings difference*. The earnings difference may be irrelevant to substantively rational agents, but we expect people to more readily learn to play a best response when the payoff from doing so is larger.

Learning-based models draw attention to two effects of variations in the earnings difference. First, many models based on deterministic dynamics, such as the replicator dynamics, predict that convergence will be faster when the earnings difference is larger. This follows directly from the fact that a larger earnings difference increases the rate at which strategies are adjusted in the direction of a best response. For the case of the continuous-time replicator dynamic, for example, the rate of convergence is twice as large in game *2R* relative to *R* and six tenths as large in game *0.6R* relative to *R*.

Secondly, the earnings difference may influence which equilibrium is selected. Binmore and Samuelson's (1997) model of aspiration and imitation modifies the Kandori, Malouf, and Rob model to accommodate a stochastic learning (as well as mutation) process. In the medium run, this opens the possibility that the noisy learning process will lead from the basin of attraction of one equilibrium to the other. Over longer periods of time, the model predicts that as the earnings difference decreases from *2R* to *R* to *0.6R* the payoff dominant equilibrium will be more likely to appear. Intuitively, the smaller payoff difference of the *0.6R* game makes it more likely that the stochastic learning process will lead away from a best response, and accordingly escape its current basin of attraction, an event we refer to as a *separatrix crossing*. The relatively low payoffs attached to the risk-dominant equilibrium of game *0.6R* interact with the players' aspirations to make it especially likely that such an escape leads from the risk-dominant to the payoff-dominant equilibrium. We accordingly expect the payoff-dominant equilibrium to be more likely to appear in game *0.6R* than in game *R* and more likely in game *R* than in game *2R*.

Our experimental results provide evidence that changing the earnings difference between *X* and *Y* influences behavior. Behavior converges faster in *2R* than in *R*, which in turn converges faster than in *0.6R*. Separatrix crossings were observed more frequently in *0.6R* than in *R*, which in turn had more separatrix crossings than *2R*. The payoff-dominant equilibrium does emerge as a convention more often in the games with a smaller earnings difference.

The following section describes the experiment. Section 3 uses various models to make predictions about behavior in the experiment. Section 4 presents the experimental results. The penultimate section relates our findings to the literature and the final section contains concluding comments.

2 EXPERIMENTAL DESIGN

Human subjects played the *2R, R* or *0.6R* stag hunt game form for seventy-five periods. Eight subjects participated in each cohort. We used a single-population random matching protocol to pair subjects.

The subjects had common and complete information about both their own and everybody else's earnings table. Subjects confronted an anonymous participant each period. Their actions were designated 1 and 2, and each subject chose one such action in each period. After strategy choices were made, the subjects were then randomly paired to determine an outcome for each pair. The subjects were informed that they were being randomly paired. Since outcomes were reported privately, subjects could not use common information about the outcomes in previous periods to coordinate on an equilibrium.

Monetary payments were used to induce preferences. Numbers in the game form denote the number of cents earned by the subjects given they chose a given action combination. Subjects were also instructed on how to derive the other participant's earnings from the earnings table.

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

The subjects were recruited from undergraduate economic classes at Texas A&M University in the spring of 1996 and fall of 1997. A total of 96 subjects participated in the experiment; four cohorts of eight subjects each playing game *2R*, four cohorts of eight subjects each playing game *R*, and four cohorts of eight subjects each playing game *0.6R*. 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. Repeated play of the payoff-dominant equilibrium for seventy-five periods, which take about two hours, results in a subject earning $33.75.

3 ANALYTICAL FRAMEWORK

3.1 THE EARNINGS DIFFERENCE

Games *2R, R,* and *0.6R* have the same best-response correspondence, but differ in the penalty attached to not playing a best response or, more optimistically, to the reward for playing a best response. We refer to this incentive as the *earnings difference,* and it is the variation in earnings differences that leads us to expect different behavior in the three games.

Let p denote the probability a player chooses X and q denote the probability his opponent chooses X. Then the expected payoff for the player is $u(p,q;M) = \{p,1\text{-}p\}.M.\{q,1\text{-}q\}$, where M denotes the relevant earnings matrix. The earnings difference, given strategy q, is given by

$$r(q;M) \equiv u(X,q;M) - u(Y,q;M) \equiv \{1,-1\}.M.\{q,1-q\}.$$

For game R, $r(q;R) = 25q\text{-}20$, for game $2R$, $r(q;2R) = 50q\text{-}40 = 2r(q;R)$, and for game $0.6R$, $r(q;0.6R) = 15q\text{-}12 = 0.6r(q;R)$. Hence, for any mixture q, the earnings difference between the two actions 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 forces attracting players to choose best responses will be more effective in games in which the earnings difference is greater.

3.2 PROBABILISTIC CHOICE

Perhaps the first choice theory to capture such a possibility was Luce's (1959) probabilistic choice model. Let A denote a set of alternatives and let a and b denote elements of A. Let $P(a,b)$ denote the probability that a person chooses a over b when making a choice from the set $\{a,b\}$ and let $P_S(a)$ denote the probability that a person chooses a when making a choice from the set $S \subseteq A$. Luce (1959, p.23) showed that if and only if behavior satisfies the independence from irrelevant alternatives, and $P(a,b) \neq 0,1$ for all a,b in A, then there exists a positive real-valued function v defined on A such that

$$P_S(a) = \frac{v(a)}{\sum_{b \in S} v(b)}.$$

The function v is unique up to multiplication by a positive constant.

We must make some assumption concerning the relationship between monetary payoffs and Luce's v-scale. If we let $v(X) = \exp(\lambda\, u(X,q))$, then we obtain the widely studied and econometrically tractable logit model

$$p(q,\lambda) \; = \; P(X,Y) \; = \; \frac{\exp(\lambda\, u(X,q))}{\exp(\lambda u(X,q)) \; + \; \exp(\lambda u(Y,q))},$$

where λ is a parameter and $p(q,\lambda)$ is the probability of X given q, and λ. Recalling the earnings difference, we can solve for the logit response function[5]

$$p(q,\lambda;M) \; = \; \frac{\exp(\lambda\, r(q;M))}{1 \; + \; \exp(\lambda\, r(q;M))}.$$

When λ=0, players are indifferent over all strategies, while setting λ equal to infinity gives best-response behavior. Fixing λ, a *logit equilibrium* is a fixed point of the two players' logit response functions (see McKelvey and Palfrey (1995)).

Following Fudenberg and Levine (1996), we can use the logit response function to define a single-population continuous-time logit response dynamic,

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

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. The stationary states of the single-population logit response dynamic correspond to symmetric logit equilibria.

Figure 4 graphs the single-population continuous-time best-response dynamic, which is the same for all three games, and the logit response dynamic. Depending on λ and $r(q;\, M)$, the logit response dynamic may or may not have asymptotically stable stationary states close to the risk-dominant equilibrium ($q=0$) and the payoff-dominant equilibrium ($q=1$). In Figure 4, λ is set to 1, so that players are responsive to payoffs but do not always choose best responses. As long as λ is neither too large nor too small, the qualitative result that a smaller earnings difference favors risk

---

[5] An alternative route to the same function is to use a random utility model, see Maddala (1983) or Anderson *et al.* (1992).

dominance will be preserved. With λ=1, both *2R* and *R* have three logit equilibria that are close to the strategic equilibria. The risk-dominant equilibrium's basin of attraction under the logit response dynamic is larger for game *R* than for game *2R*, and both are larger than under the best-response dynamic. If we think of some fixed distribution governing the initial conditions from which the dynamic begins, then the effect of probabilistic choice is to make the payoff-dominant equilibrium less likely than in the case of best-response dynamics. Similarly, given a fixed distribution, the payoff-dominant equilibrium is less likely to appear in game *R* than in game *2R*. Moreover, there is a unique logit equilibrium, corresponding to the risk-dominant equilibrium, for game *0.6R*.[6] For any initial condition, the probabilistic choice analysis with λ=1 thus predicts the emergence of the risk-dominant equilibrium for game *0.6R*.

3.3 Reinforcement Learning

The logit response dynamic is an example of the common approach of using deterministic differential equations to describe learning-based theories of equilibrium. Another commonly studied system is the replicator dynamic. Originally criticized as being of purely biological interest, it is now clear that the replicator dynamics may sometimes serve as a theoretical approximation of a learning model.[7] The replicator dynamic has also been found to provide a useful description of the relationship between initial conditions and emergent conventions in simple coordination and bargaining experiments.[8] Here we trouble the reader with the replicator dynamic because it is an easy framework within which to derive different speed-of-convergence predictions for the three games.

The basins of attraction under the replicator dynamic coincide in the three games, which in turn coincide with the basins of attraction of the myopic best-response dynamic. The replicator prediction in both games is then that a cohort with initial population frequencies of *X* less than k^* will converge to the risk-dominant equilibrium and a cohort with initial

---

[6]The uniqueness of the logit equilibrium for game *0.6R* is implied by the fact that the logit response function intersects the horizontal axis only once.

[7]For examples of choice models that are approximated by the replicator dynamics, see Binmore, Gale and Samuelson (1995), Börgers and Sarin (1997), Cabrales (1993), and Schlag (1994).

[8]See Friedman (1996), Van Huyck, Battalio, and Rankin (1997) and Van Huyck *et al.* (1995) respectively.

population frequencies of X greater than k^* will converge to the payoff-dominant equilibrium.

The replicator dynamic also makes predictions concerning the speed at which the population converges to equilibrium that do not emerge from an analysis of best-response dynamics and which are related to the earnings difference. The single-population replicator dynamic is given by the following equation:

$$\dot{q} = q(\{1,0\}.M.\{q,1-q\} - \{q,1-q\}.M.\{q,1-q\}),$$

where q is the population frequency of X and M is the earnings matrix. Figure 5 graphs the replicator dynamic for games *0.6R, R,* and *2R*, Comparing with Figure 4, we see that the best-response dynamic predicts the largest changes in behavior occur near the separatrix, while the replicator dynamic predicts that changes will be smallest near the separatrix and relatively large in the middle of the basins of attraction. It is straightforward to calculate that

$$\frac{\dot{q}}{q}\Big|_{2R} = 2\frac{\dot{q}}{q}\Big|_{R}$$
$$\frac{\dot{q}}{q}\Big|_{0.6R} = 0.6\frac{\dot{q}}{q}\Big|_{R}.$$

Hence, the replicator dynamic predicts that the rate at which the system converges to an equilibrium is twice as large in game *2R* as it is in game *R* and is six tenths as large in game *0.6R* as it is in game *R*.

From Figure 4, we can see that the logit response dynamic also makes predictions concerning speeds of convergence. However, the speed of convergence depends on the state. Sometimes the absolute value of the rate of change is largest in game *R* and sometimes the rate of change is largest in game *0.6R* under the logit response dynamic, so the logit response dynamic makes a more complicated prediction than the replicator dynamic.

3.4 Aspiration and Imitation Model

Binmore and Samuelson (1997) introduce an aspiration and imitation model in which players tend to revise their strategies whenever their payoffs fall below an aspiration level, and when doing so choose new strategies by imitating the behavior of other players. The resulting learning behavior is stochastic. Players tend to move in the direction of a best response, but occasionally abandon a best response in favor of an inferior response. As a result, the learning process can cause the proportion of the population playing strategy *X* to cross the separatrix, that is, to change from a state below k^* to a state above k^* or vice versa. These separatrix crossings are not predicted by best-response, continuous logit, or replicator dynamics.

If enough time passes, the aspiration and imitation model will lead the population of players to spend most of their time near one of the strict Nash equilibria of the stag hunt game. Proposition 8 in Binmore and Samuelson (1997, p.256) allows us to compare how the equilibrium that is selected in this way is likely to differ in the three games. In particular, altering the payoffs to action *Y* so as to reduce the earnings difference, while preserving the best-response correspondence as well as the expected payoff from the mixed equilibrium, makes a convention based on payoff dominance more likely to emerge. Hence, unlike the case of the logit response dynamic, the payoff dominant equilibrium is more likely in game *0.6R* than in game *R* and is more likely in game *R* than in *2R*.

Intuitively, the smaller payoff difference makes it more likely that players will mistakenly choose an inferior response when selecting a strategy, hence making it more likely that the system will cross the separatrix. In addition, the relatively low payoffs attached to the risk-dominant equilibrium in games with a smaller earnings difference, which results from holding earnings in the payoff dominant and mixed equilibrium constant across games, ensure that payoffs in these games are more likely to fall short of the aspiration level. Hence strategy adjustments will be more common near the risk-dominant equilibrium. Switches across the separatrix are thus more likely to lead from the risk-dominant to the payoff-dominant equilibrium than the reverse, which makes a convention based on payoff-dominance more likely to emerge.

3.5 RELATIONSHIP TO THE EXPERIMENTS

None of these theoretical learning models provides an exact description of the experimental setting in which our subjects played the game. Theoretical models which motivate the replicator dynamic as approximations of choice behavior typically describe either the behavior of single agents or the behavior of arbitrarily large populations of agents.[9] The continuous logit model is similarly derived from a model of individual choice, but then removes the randomness from the model that served as its

[9] Börgers and Sarin (1997) and Schlag (1994) fall into the former category; Binmore, Gale and Samuelson (1995) and Cabrales (1993) into the latter.

original motivation by assuming that the population is arbitrarily large.

The aspiration and imitation model of Binmore and Samuelson (1997) assumes that only one agent revises strategies in any period, with arbitrarily short time periods. In contrast, the experiment synchronizes play. The model produces equilibrium selection results for the limiting case of an arbitrarily large population. The model assumes that agents choose new strategies by imitating other randomly selected members of the population, while the experimental subjects only had information on their own history of play. The equilibrium selection prediction derived from the aspiration and imitation model is based on the properties of the model's stationary distribution, describing the limit of a process in which the model alternates between long periods of time near each of the ends of the state space, with relatively shorter periods spent traversing from one end to the other. We are not likely to observe such behavior in 75 plays of the stag hunt game.

There is room for an apologist to explain away any observation that is inconsistent with his favored model, and the experiment should not be viewed as a test of these models. However, useful models provide insight into behavior that occurs in settings which do not exactly duplicate the abstractions of the model. Our experiment can be viewed as exploring the circumstances under which these models provide insight into observable behavior. In the case of our stag hunt games, phenomena predicted by these models, such as history-contingent equilibrium selection, differences in the speed of convergence, and separatrix crossing, are observed in the experiment.

4 EXPERIMENTAL RESULTS

Section 4.1 reports the experimental results for the twelve cohorts focusing on aggregate behavior. Section 4.2 reports estimated probabilistic choice models of individual subject behavior. Section 5 summarizes the results and discusses their relationship to the literature.

4.1 SUMMARY OF AGGREGATE BEHAVIOR

All twelve cohorts start in the risk-dominant equilibrium's (*Y,Y*) basin of attraction under best-response or replicator dynamics. Specifically, for the *0.6R* cohorts, the period one frequency of action *X* is $\{\frac{4}{8}, \frac{6}{8}, \frac{6}{8}, \frac{6}{8}\}$

respectively. For the *R* cohorts, the period one frequency of action *X* is

$\{\frac{6}{8}, \frac{5}{8}, \frac{6}{8}, \frac{6}{8}\}$ respectively. For the *2R* cohorts, the period one frequency of action *X* is $\{\frac{5}{8}, \frac{4}{8}, \frac{5}{8}, \frac{3}{8}\}$ respectively. Combining the three treatments gives an average initial condition of 0.6458, or slightly more than 5 out of 8 subjects playing action *X*, the payoff dominant action. The modal subject thus plays the payoff dominant action in the first period, but not enough subjects focus on payoff dominance to make playing the payoff dominant action mutually consistent, since 0.6458 is less than $k^* = 0.8$.

The following contingency table, crossing the game and subject choice in period 1, was used to test the hypothesis that the game did not influence initial behavior. The Chi-square statistic is 2.8, which given 2 degrees of freedom has a *p-value* of 0.24. Hence, subjects' slight tendency to initially play the payoff dominant action more frequently when the earnings difference is smaller is not statistically significant at conventional levels.

Contingency Table
Treatment by Period 1 Subject Choice

	X	*Y*	Total
0.6R	22	10	32
R	23	9	32
2R	17	15	32
Total	62	34	96

Figures 6*a*,*b*,and *c* report the five-period mean frequency of the payoff-dominant action *X* by treatment. 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). All *2R* cohorts converge to the risk-dominant equilibrium and three of four *R* cohorts converge to the risk-dominant equilibrium, with the remaining *R* cohort hovering near the payoff dominant equilibrium. None of the *0.6R* cohorts converge to the risk-dominant equilibrium. Instead, two of four *0.6R* cohorts converge to the payoff dominant equilibrium, with the

remaining two cohorts hovering near the center of the state space.

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 the *2R* cohorts. If we compute the average first period in which every subject in the cohort plays the risk-dominant action (excluding the *R* cohort that never converges to the risk-dominant equilibrium), we find that the remaining three *R* cohorts take an average of 50 periods for all agents to reach the risk-dominant equilibrium, while the four *2R* cohorts take an average of 29 periods. The evidence is thus consistent with the prediction that reducing the earnings difference reduces the speed of convergence.

<insert frequency figures about here>

In the time averaged data, there are only four separatrix crossings (see Figure 6 *a,b,c)*. This is out of a potential of 168, so separatrix crossings are rare.[10] Two *0.6R* cohorts account for two separatrix crossings and one *R* cohort accounts for two. The two crossings in the *0.6R* treatment and one of the crossings in the *R* treatment are of the expected type, that is, the crossing is from the risk-dominant equilibrium's basin of attraction to the payoff-dominant equilibrium's basin of attraction. The *2R* cohorts never crossed the separatrix in the time averaged data.

For the *0.6R* cohorts, the period 75 frequency of action *X* is $\{\frac{3}{8}, \frac{8}{8}, \frac{8}{8}, \frac{2}{8}\}$ respectively. For the *R* cohorts, the period 75 frequency of action *X* is $\{\frac{1}{8}, \frac{0}{8}, \frac{5}{8}, \frac{0}{8}\}$ respectively. For the *2R* cohorts, the period 75 frequency of action *X* is $\{\frac{1}{8}, \frac{0}{8}, \frac{0}{8}, \frac{0}{8}\}$ respectively. Our results are thus consistent with the proposition that game *0.6R* is less likely to converge to the risk-dominant equilibrium than is game *2R.*

These results reflect the qualitative features predicted by the replicator

[10]There are 32 (out of a possible 888) separatrix crossings in the raw data. However, most of these crossings are reversed immediately. For example, the *2R* cohorts account for 8 separatrix crossings in the raw data of which 4 are crossings from the risk dominant to the payoff dominant basin of attraction. The other four crossings occurred immediately after a risk dominant to payoff dominant crossing and are crossings from the payoff dominant basin back to the risk dominant basin.

dynamic and the aspiration and imitation model discussed in section 3. As predicted by the replicator dynamic, convergence is more rapid when the earnings difference is larger. The risk-dominant equilibrium emerges as the customary way to play in all of the *2R* cohorts and in three out of four *R* cohorts. The risk-dominant equilibrium never emerged in the four *0.6R* cohorts. Conversely, the payoff-dominant equilibrium emerges as the customary way to play in two *0.6R* cohorts. This ability of populations to escape the risk-dominant basin of attraction when the earnings difference is low is consistent with the aspiration and imitation model. Separatrix crossings occur but are rare, and are more likely when the earnings difference is small.

4.2 PROBABILISTIC CHOICE MODELS OF INDIVIDUAL BEHAVIOR

In this section, we report estimates of probabilistic choice learning models based on the individual subject data. These provide some insight into how the varying earnings differences of games *2R*, *R*, and *0.6R* are reflected in individual behavior.

Popular belief-based learning models consist of two components: an assessment rule used to describe the opponents' behavior and a response function. For example, fictitious play consists of a best response function and an assessment rule based on the historical frequency of opponents' actions. Exponential fictitious play retains the assessment based on historical frequencies but replaces the best-response function with the logistic response function.

Fictitious Play Beliefs

Let y_{it} denote subject *i*'s assessment of the likelihood his opponent will play action *X* in period *t*. For a history of observed actions given by h_{it} = (s_{i1}, s_{i2}, ...,s_{it-1}), fictitious play and exponential fictitious play specify y_{it} as

$$y_{it} = \frac{t-1}{t} y_{it-1} + \frac{1}{t} s_{it-1}, \qquad t \geq 2,$$

where s_{it} is 1 if subject *i*'s opponent chose action *X* in period *t* and zero otherwise. Let y_{i1} equal k^*.

We estimated the following generalized model of exponential fictitious play (EFP):

$$p_{it} = f(\alpha_i + \beta_i r(y_{it})),$$

where $f(\bullet)=exp(\bullet)/(1+exp(\bullet))$ is the logistic response function and $r(\bullet)$ is the earnings difference. When α_i is zero, it is straightforward to show that this is Fudenberg and Levine's (1996, p.153) model of exponential fictitious play.

The estimated EFP model breaks the parameters α_i, and β_i into a representative component α and β, and an idiosyncratic (individual subject) component $\hat{\alpha}_i, \hat{\beta}_i$, that is, $\alpha = \alpha_i - \hat{\alpha}_i$ and $\beta = \beta_i - \hat{\beta}_i$. The model was estimated using the logistic procedure in *SAS* version 6.12. Variables were chosen for inclusion in the model using the forward selection option of the logistic procedure, which adds variables iteratively according to the score chi-square statistic until there are no variables that pass the five percent statistical significance rule of thumb.

Table 1 reports four probabilistic choice learning models that use fictitious play beliefs. The table reports the representative component of the estimated model for all 96 subjects combined, where *std* denotes standard error, *trt1* is a treatment dummy variable for the *2R* sessions, *trt2* is a treatment dummy variable for the *0.6R* sessions, *n* denotes number of observations, *df* denotes degrees of freedom, χ^2 denotes the chi-square statistic for the global hypothesis that all βs are zero, *H/L* is the Hosmer/Lemeshow goodness-of-fit measure, and *p-value* denotes the probability value of the *H/L* statistic. The table does not report the idiosyncratic components of the fitted model, but by subtracting the number of representative components included from *df*, one can obtain the number of idiosyncratic components.

Model 1 in Table 1 is a generalized version of fictitious play ($p_{it}=f(\alpha_i+\beta_i\, y_{it})$). It is reported for comparison with exponential fictitious play. As we would expect, the estimated coefficient of the belief variable y_{it} is positive (equaling 9.22), indicating that subjects were more likely to play strategy *X* the more they expected opponents to play *X*. Both treatment dummy variables are significant, indicating that behavior differed between the three treatments.

Model 2 is exponential fictitious play. The belief variable, y_{it}, is now transformed by the earnings difference function, $r(y_{it})$. Note that this transformation changes units from frequency to cents. Hence, the magnitude of the parameter estimates are not directly comparable. The treatment dummy variables are no longer included by the procedure. The difference in behavior between the games can thus be summarized by the earnings difference variable and its effect on probabilistic choice.

Table 1: Estimated Fictitious Play and Exponential Fictitious Play Learning Models: Periods 1 to 40.

Model	α (*std*)	y_{it} (*std*)	$r(y_{it})$ (*std*)	*trt1* (*std*)	*trt2* (*std*)	*n* (*df*)	-2logL (χ^2)	H/L *p-value*
1	-4.772	9.22	n/a	-3.12	-1.25	3840	3007.3	10.13
All Cohorts	(0.23)	(0.38)		(0.25)	(0.17)	(67)	(2240)	0.26
2	1.856	n/a	0.348	n/i	n/i	3840	3021.4	21.81
All Cohorts	(0.09)		(0.02)			(59)	(2225)	0.00
3	1.328	n/a	0.097	n/i	n/i	1280	1279.1	40.70
0.6R Cohorts	(0.11)		(0.03)			(18)	(447)	0.00
4	2.248	n/a	0.408	n/i	n/i	2560	1701.7	8.58
R&2R Cohorts	(0.15)		(0.02)			(41)	(1597)	0.38

n/a - not allowed; n/i - allowed but not included.

Figure 7 is a graph of the response dynamic based on model 2. This can be compared with the continuous logit dynamic of Figure 4. The main difference between Figures 4 and 7 is the estimated value α =1.856 (see model 2 in table 1), which is statistically significant at less than the one percent level. If α had been 0 and β had been 1, then the figures would have been the same. The positive value of α causes the basin of attraction of the payoff-dominant equilibrium to be larger than would be the case under the continuous logit dynamic. Moreover, this effect is larger for treatments with smaller earnings differences because they have shallower basins of attraction.

The combination of the payoff dominant equilibrium's exogenous attraction and the relative shallowness of game *R's* basin of attraction results in game *R* having a payoff dominant equilibrium with a larger payoff basin of attraction than game *2R*. For game *0.6R*, the combination of the payoff dominant equilibrium's exogenous attraction and the shallowness of the basin completely eliminates the risk dominant equilibrium's basin of attraction. The only logistic equilibrium is close to the payoff dominant equilibrium. The estimated model also predicts a region around 0.25 in which the dynamic is close to being stationary. The estimated exogenous attraction of the payoff dominant action thus completely reverses the prediction of the continuous logit dynamic with α = 0.

The EFP model fits our data surprisingly well. Our finding of an exogenous attraction to the payoff dominant action contrasts with Van Huyck, Battalio, and Rankin (1997), who found α to be 0 in a coordination game without Pareto ranked strict equilibria. Our finding also contrasts with Mookerjee and Sopher (1994), which can be attributed to the differences between a random and a repeated-pairs matching protocol. An inertial dynamic, like fictitious play, is much more difficult to exploit in a random-matching protocol than in a repeated-pairs protocol.[11]

While the earnings difference captures an important feature of the data, the second model has a slightly higher likelihood score than the first. Moreover, the Hosmer/Lemeshow goodness-of-fit test, H/L, rejects the second model. Both of these features are due to including the *0.6R* cohorts in the data used to estimate the model. Models 3 and 4 report estimates of the EFP model for the *0.6R* cohorts and for the *R* and *2R* cohorts respectively. The bad fit for the *0.6R* cohorts arises because the EFP model, with all twelve cohorts included, predicts a stable, interior equilibrium when both the estimated β and the earnings difference is small. The model thus predicts that, as the belief that *X* is the customary way to play goes to 1, subjects will *reduce* the likelihood they will choose the payoff-dominant action.[12] However, the cohorts in which subjects are observing y_{it} close to 1 converge quickly and surely to perfect conformity with payoff dominance, that is, all eight subjects play the payoff dominant action. Hence, the model fits behavior in these cohorts badly. When we estimate the model excluding the *0.6R* cohorts we obtain a model with essentially the same qualitative features as the estimated model for all twelve cohorts, and one that passes the Hosmer/Lemeshow goodness-of-fit test.

Reinforcement Learning

The previous section provided evidence that behavior responds to the expected earnings difference, which is based on beliefs about opponents' play gleaned from previous experience. A number of investigators have noted that the expected earnings difference can be approximated by the average earnings difference experienced in previous play. This approach may be a cognitively simpler procedure because one does not have to

---

[11]This distinction between protocols that make it easy to exploit adaptive behavior and those that don't is demonstrated nicely in Bloomfield (1994). The repeated minimum rule game, which Camerer and Ho (1996) use to reject both reinforcement and belief based learning, seems intuitively to be between repeated and evolutionary models since only the minimum matters. Sefton (1996), studying an evolutionary coordination game similar to ours, gets similar results.

[12]Figure 7 reveals this feature: note how the rate of change is predicted to be negative at y_{it} close to 1 for *0.6R*.

transform the frequency of opponent's actions into expected payoffs (see, for example, Erev and Roth (1995)). We refer to such a procedure as reinforcement learning because actions depend only on the differences in the earnings actually experienced by the players.

Let a_{it} equal 1 if subject i chose action X in period t and 0 otherwise. Recall that s_{it} equals 1 if subject i's opponent chose action X in period t and zero otherwise. Let $u(a_{it}, s_{it})$ denote subject i's earnings in period t given a_{it} and s_{it}. Let subject i's average earnings through period t when choosing X, $u_{it}(X)$, and subject i's average earnings through period t when choosing Y, $u_{it}(Y)$, be given by

$$u_{it}(X) = \frac{(0.36 + \sum_{\tau=1}^{t} a_{i\tau}\, u(a_{i\tau}, s_{i\tau}))}{1 + \sum_{\tau=1}^{t} a_{it}};$$

$$u_{it}(Y) = \frac{(0.36 + \sum_{\tau=1}^{t} (1-a_{i\tau})\, u(a_{i\tau}, s_{i\tau})}{1 + \sum_{\tau=1}^{t} (1-a_{it})}.$$

Notice that we initialize the reinforcement algorithm with the expected earnings from the mixed strategy equilibrium 0.36, which corresponds to the expected earnings given a prior belief that X will be played with probability 0.8. Then subject i's experienced earnings difference between action X and Y through period t is given by

$$r_{it} = u_{it-1}(X) - u_{it-1}(Y).$$

Table 2 reports two versions of the reinforcement learning model $p_{it}=f(\alpha_i+\beta_{1i}\, r_{it}+\beta_{2i}\, r(y_{it}))$. Model 5 estimates a pure reinforcement learning version that excludes the expected earnings difference. The experienced earnings difference variable, r_{it}, is statistically and economically significant. The treatment dummy is also statistically and economically significant. The value of the likelihood score is larger than the belief-based models and the Hosmer/Lemeshow goodness-of-fit statistic rejects the model.

Table 2: Estimated Reinforcement Learning Models: Periods 1 to 40.

Model	α (*std*)	r_{it} (*std*)	$r(y_{it})$ (*std*)	*trt1* (*std*)	*trt2* (*std*)	*n* (*df*)	-2logL (χ^2)	H/L *p-value*
5	1.031	0.173	na	-1.322	n/i	3840	3565.7	20.17
All Cohorts	(0.07)	(0.01)		(0.12)		(61)	(1681)	(0.01)
6	2.963	0.048	0.422	-1.742	-1.810	3840	2865.7	30.02
All Cohorts	(0.19)	(0.01)	(0.02)	(0.26)	(0.19)	(71)	(2381)	(0.00)

na - not allowed; n/i - not included.

Model 6 includes both the experienced earnings difference variable, r_{it}, and the expected earnings difference variable, $r(y_{it})$. Surprisingly, the forward procedure includes both variables. Including the expected earnings difference reduces the parameter for the experienced earnings difference by more than half, while the magnitude of the parameter for the expected earnings difference variable is larger than is the case without the experienced earnings difference (see Table 1). A change in the expected earnings difference has about four and one half times as large an impact when compared to a change in the experienced earnings difference. Both treatment dummy variables are included. The model gives the smallest likelihood score of those reported above, but the Hosmer/Lemeshow goodness-of-fit test rejects the model.

This result is similar to Camerer and Ho's (1996) finding that an "Experience-weighted Attraction" (EWA) model, including both experienced earnings and expected earnings, fits their data better than either a pure reinforcement learning or belief-based learning algorithm. An important difference between the EWA model and our empirical models is that we allow idiosyncratic subject behavior.[13] All of our models have statistically and economically significant individual-subject intercept and slope dummy variables.

Models with Lagged Choices

In a rather different context, a repeated matching pennies game, Mookerjee and Sopher (1994) report evidence of significant own and other lagged choices in models of reinforcement and belief-based learning. Specifically, they found evidence of negatively autocorrelated choices, which is

[13]Chueng and Friedman (1996) and Stahl and Wilson (1995) also find evidence for idiosyncratic behavior.

inconsistent with repeated play of the unique mixed strategy equilibrium of the stage game. Finding evidence of autocorrelated behavior in evolutionary stag hunt games should not be surprising, since this pattern results from an emergent convention. Here we extend the models from the previous sections to allow own and other lagged choices.

Table 3 reports exponential fictitious play and combination models with lagged choice variables included. The lagged own, *scl1*, and other, *pcl1*, choice variables are always significant. Both playing *X* and observing *X* last period increases the likelihood a subject will play *X* in the current period.

Table 3: Probabilistic choice learning models with lagged own and other choice: Periods 2 to 40.

Model	α (*std*)	$r(y_{it})$ (*std*)	r_{it} (*std*)	*trt1* (*std*)	*trt2* (*std*)	*scl1* (*std*)	*pcl1* (*std*)	*n* (*df*)	-2logL (χ^2)	H/L *p-value*
7 *All Cohorts*	n/i	0.211 (0.09)	n/a	n/i	n/i	1.441 (0.08)	1.017 (0.08)	3744 (44)	2660.7 (2529)	23.40 (0.00)
8 *All Cohorts*	n/i	0.237 (0.01)	0.048 (0.01)	n/i	-0.649 (0.12)	1.446 (0.09)	1.203 (0.10)	3744 (57)	2566.6 (2623)	20.84 (0.01)
9 *0.6R Cohorts*	-1.217 (0.150	n/i	n/i	n/a	n/a	2.215 (0.16)	1.133 (0.16)	1248 (13)	1073.4 (609)	12.9 (0.07)
10 *R&2R Cohorts*	2.763 (0.39)	0.402 (0.04)	0.095 (0.02)	-2.90 (0.37)	n/a	0.689 (0.15)	1.062 (0.16)	2496 (54)	1421.6 (1770)	2.41 (0.97)

n/a - not allowed; n/i - not included.

Model 7 is the EFP model (Model 2) with lagged choices. The estimated model no longer includes a significant positive intercept and the magnitude of the parameter for the earnings difference variable is halved (compare with Table 1).

Model 8 in Table 3 combines the expected and experienced earnings variable (Model 6) with lagged choices. The intercept is not included. The magnitude of the treatment dummy and expected and experienced earnings parameters are reduced but have the same qualitative characteristics as in Table 2.

In attempting to solve the goodness-of-fit problem reflected in the Hosmer/Lemeshow values of models 7 and 8, we estimated a number of models that altered the belief variable either by optimizing the priors, allowing for treatment specific priors, or by introducing Cheung and Friedman's (1996) memory discounting, which both eliminates the prior and introduces memory discounting. None of these attempts solved the goodness-of-fit problem.

Model 9 restricts the data to *0.6R* cohorts. Neither the expected or experienced earnings difference is included by the procedure. There are 11 idiosyncratic components. Two subjects have an idiosyncratic slope parameter. Two subjects have an idiosyncratic experienced earnings parameter. Five subjects have an idiosyncratic expected earnings parameter. One subject has both an idiosyncratic experienced and expected earnings parameter. So while the procedure no longer reports a significant influence of the earnings difference variables for the representative subject there are individual subjects who are influenced by the earnings difference variables. Overall, however, the earnings difference models do not fit the *0.6R* cohorts data very well once one accounts for the correlation in own and other choices.

Model 10 restricts the data to the *R* and *2R* cohorts. Own and other lagged choices are again positively correlated with the response variable. Both the expected and experienced earnings differences are significant, but the magnitude of the expected earnings difference parameter is more than four times as large as the magnitude of the experienced earnings difference parameter. The positive intercept indicates an unexplained preference for the payoff dominant action, *X*. Model 10 passes the Hosmer/Lemeshow goodness-of-fit test.

5 RELATED LITERATURE AND DISCUSSION

Table 4 summarizes our results and compares them to previous findings. Each row represents a cohort. The cohorts are ordered first by value of the separatrix, k^*, so that the basin of attraction of the risk-dominant equilibrium shrinks as one moves down the table, and second by the size of the scaled earnings difference, $R(q)$, that is, the earnings difference divided by the game's highest possible payoff. The initial and terminal outcomes are reported as the ratio of subjects using the payoff-dominant action to the total number of subjects active in the cohort. The last two columns report the number of periods and the source. The experiments differ in many details such as matching protocol, induced value technique, and the cohort's experience as a group with pretrial games. We don't focus on these differences because we think the results in the literature tell a fairly consistent story.

First, subjects do not bring risk dominance into the laboratory. The payoff dominant action is usually the modal initial choice even when k^* takes on extreme values. Changing the scaled earnings difference has little discernable influence on initial conditions.

Second, the experimental subjects typically approach a mutual best response outcome, that is, the cohort converges to a customary way to solve their strategy coordination problem that is based on their experience with the cohort. The emergent convention is usually the inefficient, risk-dominant equilibrium when $k^* > 0.75$ and the efficient, payoff-dominant equilibrium when $k^* < 0.5$. For $0.5 < k^* < 0.75$, results are mixed.

Third, in most cases the terminal outcome is accurately predicted by the location of the initial outcome in the respective equilibrium's basin of attraction. Separatrix crossings occur, but are rare.

Finally, the earnings difference between the two actions influences the frequency of observed separatrix crossings. Our experiment provides an explicit treatment of the earnings difference, while circumstantial evidence appears in two other cases.

Straub (1995) found that play converged to the risk-dominant equilibrium in cohorts 10 and 19, but that play converged to the payoff-dominant equilibrium in cohort 26.[14] Of these three cohorts, cohort 26 played a stag game that has the smallest earnings difference. The outcomes of these three cohorts are then consistent with our finding, and the prediction of Binmore and Samuelson's (1997) aspiration and imitation model, that the payoff-dominant equilibrium will be most likely to appear when the earnings difference is small. However, the basin of attraction of the payoff-dominant equilibrium is also the largest in the case of cohort 26. Cohorts 10, 19, and 26 played games having values of k^* equal to 0.8, 0.75, and 0.67 respectively, so there is a reason for the payoff-dominant equilibrium to appear in cohort 26, but not cohorts 10 and 19, that is not related to the earnings difference.

Friedman (1996) does not report his results by cohorts or in some cases even by game forms. We thank him for making his raw data available so that we can make comparisons with our experiment. Table 4 reports data for the first encounter a cohort had with a stag hunt game in the long sequence of game forms and other treatment changes experienced by that cohort. These data do not indicate that the game form influences behavior.

[14]Straub's experiments involved groups of ten subjects who each played nine repetitions of the game, one against each possible opponent.

Specifically, cohorts 21 to 25 begin and end in the basin of attraction of the risk dominant equilibrium. Cohort 20, which is the only naive cohort to play the game with double the payoff difference, also begins and ends in the risk dominant equilibrium's basin of attraction. Evidence from Friedman's (1996) naive cohorts thus does not support the Binmore/Samuelson prediction that the earnings difference will influence equilibrium selection.

However, Friedman (1996) does report finding behavioral changes consistent with our expectations for games forms run latter in the session that hold the separatrix constant and change the earnings difference. Sorting out the reasons for this difference is complicated because his experimental design changes action labels, game forms, matching rules, group size, and information within a session. He concludes that "one can bias convergence towards the other 'payoff dominant' evolutionary equilibrium."

Constantly changing action labels and game forms forces subjects to focus on deductive selection principles, like payoff dominance or risk dominance. Rankin, Van Huyck, and Battalio (1997) report an experiment in which payoff dominance emerges as a deductive selection principle in sequences of similar but not identical stag hunt games. This is so even when the stage game has an extreme value for k^*, such as 0.97. So in some cohorts players can become very confident in the mutual salience of payoff dominance.

6 CONCLUSION

In this paper, we have reported a controlled experiment that focuses on the earnings difference as the reason for separatrix crossings and the emergence of a convention based on payoff dominance. Our results provide evidence that more than the best-response correspondence matters when predicting human behavior in laboratory experiments. We focused on 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 have different pecuniary incentives to play a best response. A number of analytical models, including probabilistic choice models, deterministic replicator dynamics, and a stochastic model of aspiration and imitation, predict that the earnings difference will influence behavior. We find statistically and economically significant evidence that the expected earnings difference function helps explain observed behavior.

Our finding that convergence to an equilibrium occurs more quickly in games with a larger earnings difference is consistent with models of adaptive learning. The models of individual subject behavior estimated in Section 4.2 find some evidence of adaptive learning. From the view point of the probabilistic choice model developed in Section 3.1, in contrast, there are some anomalies. Our estimated basins of attraction are different from the predictions of the best-response or replicator dynamic, but not in the way predicted by the analysis of logit equilibria. We think this is because the abstraction assumption used to map money payoffs into the probability of actions in standard probabilistic-choice specifications does not capture the influence of the differing sizes of the payoff dominance relation between the two treatments, that is, earnings in the risk dominant equilibrium are twice as large in *2R* than in *R*.

The influence of this payoff dominance relation provides evidence for an aspiration-based model of adaptive behavior. The observed correlation between the earnings difference and the frequency of separatrix crossings is consistent with a model like Binmore and Samuelson (1997). Further work is required to more carefully investigate the difference between this and a host of competing explanations.

REFERENCES

L. Anderlini, "Communication, computability, and common interest games," Economic Theory Discussion Paper 159, St. John's College, Cambridge, 1990.

Simon P. Anderson, André de Palma, and Jacques-François Thisse, *Discrete Choice Theory of Product Differentiation*, (Cambridge, MA: The MIT Press, 1992).

Ken Binmore, John Gale, and Larry Samuelson, "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior* 8, 1995, 56-90.

Ken Binmore and Larry Samuelson, "Muddling through: Noisy equilibrium selection" *Journal of Economic Theory* 74(2) June 1997, 235-265.

Robert Bloomfield, "Learning a mixed strategy equilibrium in the laboratory," *Journal of Economic Behavior and Organization*, 25, 1994, 411-36

Tilman Borgers and Rajiv Savin, "Learning through Reinforcement and the Replicator Dynamics," *Journal of Economic Theory* 77, 1997, 1-14.

Antonio Cabrales, "Stochastic Replicator Dynamics" Mimeo, Universitat Pompeu Fabra, Barcelona, 1993.

Colin Camerer and Teck-Hua Ho, "Experience-weighted Attraction Learning in Games: Estimates from Weak-Link Games," laser-script November 1996.

Hans Carlsson and Eric van Damme, "Global games and equilibrium selection," *Econometrica*, 61: 989-1018, 1993.

Yin-Wong Cheung and Daniel Friedman, "Individual Learning in Normal Form Games: Some Laboratory Results," laser-script December 1995.

Kenneth Clark, Stephen Kay and Martin Sefton, "When are Nash Equilibria Self-Enforcing? An Experimental Analysis," laser-script, May 1996.

Russell Cooper, Douglas V. DeJong, Robert Forsythe, and Thomas W.

Ross, “Communication in coordination games,” *Quarterly Journal of Economics*, 107:739-773, 1992.

Erev, Ido and Alvin E. Roth, “On the need for low rationality, cognitive game theory: Reinforcement learning in experimental games with unique, mixed strategy equilibria”, laser-script August 1995.

Daniel Friedman, “Equilibrium in evolutionary games: Some experimental results,” *Economic Journal*, 106:1-25, 1996.

Drew Fudenberg and David K. Levine, *Theory of Learning in Games*, laser-script December 1996.

John C. Harsanyi, “A new theory of equilibrium selection for games with complete information,” *Games and Economic Behavior*, 8:91-122, 1995.

John C. Harsanyi and Reinhard Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, Massachusetts, 1988.

John Hillas, “On the definition of the strategic stability of equilibria,” *Econometrica*, 58:1365-1390, 1990.

Michihiro Kandori, George J. Mailath, and Rafael Rob, “Learning, mutation, and long run equilibria in games,” *Econometrica*, 61:29-56, 1993.

Elon Kohlberg and Jean-Francois Mertens, “On the strategic stability of equilibria,” *Econometrica*, 54:1003-1038, 1986.

R. Duncan Luce, *Individual Choice Behavior: A Theoretical Analysis* (New York,NY: John Wiley & Sons, 1959).

G.S. Maddala, *Limited-Dependent and Qualitative Variables in Econometrics*, (New York, NY: Cambridge University Press, 1983)

Richard D. McKelvey and Thomas R. Palfrey, “Quantal Response Equilibria for Normal Form Games,” *Games and Economic Behavior*, 10(1) July 1995, 6-38.

Dilip Mookherjee and Barry Sopher, “Learning Behavior in an

Experimental Matching Pennies Game," *Games and Economic Behavior* 7(1), July 1994, 62-91.

Frederick Rankin, John Van Huyck, and Raymond Battalio, "Strategic Similarity And Emergent Conventions: Evidence from Scrambled Payoff Perturbed Stag Hunt Games", laser-script, July 1997.

Arthur J. Robson and Fernando Vega-Redondo, "Efficient equilibrium selection in evolutionary games with random matching," *Journal of Economic Theory* 70 (1996), 65-92.

Karl Schlag, "Why Imitate, and if so, How?," SFB Discussion Paper B-243, University of Bonn, 1994.

Martin Sefton, "Modelling Behavior in Coordination Game Experiments," Laser-script, October 10, 1996.

Dale O. Stahl and and Paul W. Wilson, "On players' models of other players: Theory and Experimental Evidence," *Games and Economic Behavior* 10(1), 1995, 218-254.

Paul Straub, "Risk Dominance and Coordination Failure in Static Games," *The Quarterly Review of Economics and Finance*, 35(4) Winter 1995, 339-363.

Eric van Damme, "Refinements of Nash equilibrium," In Jean Jacques Laffont, editor, *Advances in Economic Theory: Sixth World Congress*. Cambridge University Press, Cambridge, 1992.

J.B. Van Huyck, R.C. Battalio, S. Mathur, A. Ortmann and P.P. Van Huyck, "On the Origin of Convention: Evidence from symmetric bargaining games," *International Journal of Game Theory* 24(2) 1995, 187-212.

John B. Van Huyck, Raymond C. Battalio, and Frederick Rankin, "On the Origin of Convention: Evidence from Coordination Games," *Economic Journal* 107(442) May 1997, 576-597.

John Van Huyck, Joe Cook, and Raymond Battalio, "Adaptive Behavior and Coordination Failure," forthcoming in *Journal of Economic Behavior and Organization.*

Jürgen Weibull. *Evolutionary Game Theory*. MIT Press, Cambridge, 1995.

Peyton Young, “The evolution of conventions,” *Econometrica*, 61: 57-84, 1993.

APPENDIX A:

Instructions text file for graphical user interface.

Doesn't include text markup symbols, page breaks, or graphics.

INSTRUCTIONS

This is an experiment in the economics of strategic decision making. Various agencies have provided funds for this research. If you follow the instructions and make appropriate decisions, you can earn an appreciable amount of money. At the end of today's session, you will be paid in private and in cash.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your cooperation.

You will be making choices on a Logitech mouse, which is located on the mouse pad in the middle of your table. You may move the pad to the right or left if this would be more comfortable. Hold the mouse in a relaxed manner with your thumb and little finger on either side of the mouse. Rest your wrist naturally on the table surface. When you move the mouse, let your hand pivot from the wrist. Use a light touch. Your cursor (a white arrow on your screen) should move when you slide the mouse on the mouse pad. If it does not, raise your hand.

To participate, you must be able to move the cursor onto an object and click any one of the mouse buttons. We will call pointing at an object and then clicking your mouse "clicking on" an object displayed on the screen. Click on the page down icon located below to display the next page.

The experiment consists of seventy-five separate decision making periods. In this experiment you will participate in a group of eight people. At the beginning of period one, each of the participants in this room will be randomly assigned to a group of size eight and will remain in the same group for the entire seventy-five decision making periods of the experiment. Hence, you will remain grouped with the same seven other participants for the next seventy-five periods.

At the beginning of each decision making period you will be randomly re-paired with another participant in your group. Hence, at the beginning of each decision making period, you will have a one in seven chance of being matched with any one of the other seven participants from your group.

At the beginning of each period, you and all other participants will choose an action. An earnings table is provided which tells you the earnings you receive given the action you and your currently paired participant chose. The actions you may choose are row 1 or row 2. Everyone has the same earnings table which will be continuously displayed on the monitor in the front of the room during the experiment. Click on the page down icon located below to display the next page.

Your earnings each period will be found in the box determined by your action and the action of the participant that you are paired with for the current decision making period. Your action determines the row and the other participant's action determines the column of the earnings table. The value in the box determined by the intersection of the row and column chosen is the amount of money that you earn in the current period. The earnings, displayed in green, in each earnings cell is the amount of money, in cents, that you earn.

The earnings each period for the participant that you are currently paired with can be determined by reversing your positions.

Click on the page down icon now to view the earnings table while I describe how the earnings for each decision making period are calculated. You can review this page at any time during the experiment by returning to the instructions.

<EARNINGS TABLE GOES HERE>

MAIN SCREEN

We will now view the main screen. You will use the main screen to make your choices each period. While you view the main screen I will read the description of the screens contained in the next two pages. You can review the text that I am reading at any time during the experiment by returning to the instructions. Click on the word "MAIN" located on the second line down from the top of the screen now. (The second line is the light blue line on your screen).

The top line of the main screen displays the title of the screen and the current period number. The second line has word "PROCEED" the abbreviation "INSTR" and the word "RECORD" on it. During the session you will be able to return to these instructions by clicking on "INSTR." You will also be able to view the history of play by clicking on "RECORD", which we will explain in a moment. The remainder of the screen is devoted to the earnings table.

Please look at the monitor at the front of the room while I demonstrate how you make and enter a choice. Do not use your main screen until you are instructed to do so.

Making a choice consists of clicking any mouse button while the cursor is in the row of your choice. When you have clicked on the earnings table, your cursor is replaced by a green highlight around the row that contained the cursor when you clicked the mouse. You can change the highlighted row by sliding your mouse up or down. Click the mouse a second time and your cursor returns, but a row remains highlighted. To enter your choice for the current period you need to confirm your choice. You confirm your choice by first clicking on the word "PROCEED" and then clicking on "YES" to confirm and enter your choice for the current period. This confirmation step lets you catch any mistakes you make.

Please make a choice now, click on proceed and then click on "NO". Notice that the row is no longer highlighted and you may now make a different choice.

Before making another choice click on "PROCEED" without making a choice and notice that you receive the following message:

YOU MUST MAKE A CHOICE BEFORE PROCEEDING

At the time this message is present, a red box is also pulsing around the outside of your earnings table.

Please make a choice now, click on "PROCEED" and then confirm your choice by clicking on "YES".

WAITING SCREEN

During a session a waiting screen will appear after you have made a choice. While you are waiting, you can view the instructions and the record of play by clicking on "INSTR" or "RECORD." When all participants have made a choice for the current period you will be automatically switched to the outcome screen. The choice displayed is the choice that you made during the demonstration of the main screen. You will automatically return to the instructions. Click on "WAITING" now.

OUTCOME SCREEN

During a session, after everyone has made their choices, the outcome screen will appear. The outcome screen summarizes what happened each period for ten seconds. Your choice will be highlighted in green. The column determined by the other participant's choice will be highlighted in . The screen is not active.
The choice displayed for your choice reflects the choice you made during the demonstration of the main screen. You will automatically return to the instructions. Click on "OUTCOME" now.

RECORD SCREEN

The record screen records the period outcomes and updates your earnings balance. A copy of the record screen is given at the top of this screen. The first three entries on the record screen are: "Period", "Your Choice" and "Other Participant's Choice ". The record screen will indicate your choice in green each period. The fourth entry is your earnings for a period which are recorded under the entry "Your Earnings". Finally, your current balance, which includes all of your earnings up to and including the current period, will be recorded under the entry "Balance". In the first period your balance is zero.

During the session the record screen will be displayed for twenty seconds. You may proceed to the next period by clicking on "RETURN" before the twenty seconds have expired. Remember you can always return to the record screen from either the main screen or the waiting screen.

Click on the word "RECORD" located on the second line down from the top of your screen now. As the experiment proceeds the records for the earlier periods will scroll off the top of the record screen. You may review the earlier records by clicking on the page up, page down, line up and line down icons located at the bottom of the record screen. Click on "RETURN" now to return to the instructions before twenty seconds have expired.

QUESTIONNAIRE

We will now pass out a questionnaire to make sure that all participants understand how to read the earnings table. Please fill it out now. Raise your hand when you are finished and we will collect it. If there are any mistakes on any questionnaire, I will go over the relevant part of the instructions again. Do not put your name on the questionnaire.

Click on the page down icon located below to display the next page.

SUMMARY

*** At the beginning of period one, each of the participants in this room will be randomly assigned to a group of size eight and will remain in the same group for the entire seventy-five decision making periods of the experiment.
*** Each period you will be randomly re-paired with one of the seven other participants in your group. Hence, at the beginning of each decision making period, you will have a one in seven chance of being matched with any one of the seven other participants your group.
*** You make a choice by clicking on a row, which highlights the row in green; clicking the mouse a second time, which restores your cursor, and then clicking on proceed and yes to confirm your choice of the highlighted row.
*** Remember that you can view the instructions or the record screen by clicking on the

appropriate word on the light blue bar.
*** Remember that you may proceed to the next period by clicking on "RETURN" before the twenty seconds have expired. You can always return to the record screen from either the main screen or the waiting screen.
*** Your balance at the end of the session will be paid to you in private and in cash.

Click on the page down icon located below to display the next page.

We have completed the instructions. Again, it is important that you remain silent and do not look at other people's work.

If you have a question, please raise your hand, and an experimenter will come to assist you. If there are no questions, period one of the experiment will begin.

Table 4: Recent Evidence on Human Behavior in Evolutionary Stag Hunt Games.

N	Game = {A, C} {B, D}	*R(k)*	*k**	Initial Outcome	Terminal Outcome	Periods	Source
1	{45,0},{35,40}	(10*k* - 8)/9	0.80	5/8	1/8	75	this paper
2	{45,0},{35,40}	(10*k* - 8)/9	0.80	4/8	0/8	75	this paper
3	{45,0},{35,40}	(10*k* - 8)/9	0.80	5/8	0/8	75	this paper
4	{45,0},{35,40}	(10*k* - 8)/9	0.80	3/8	0/8	75	this paper
5	{100,0},{80,80}^{T}	(5*k* - 4)/5	0.80	6/10	0/10	22	Cooper, *et al.* (1992)
6	{100,0},{80,80}^{T}	(5*k* - 4)/5	0.80	6/10	0/10	22	Cooper, *et al.* (1992)
7	{100,0},{80,80}^{T}	(5*k* - 4)/5	0.80	5/10	1/10	22	Cooper, *et al.* (1992)
8	{100,0},{80,80}^{T}	(5*k* - 4)/5	0.80	7/20	4/20	10	Clark, *et al.* (1996)
9	{100,0},{80,80}^{T}	(5*k* - 4)/5	0.80	5/20	2/20	10	Clark, *et al.* (1996)
10	{100,0},{80,80}^{T}	(5*k* - 4)/5	0.80	4/10	0/10	9	Straub (1995)
11	{45,0},{40,20}	(5*k* - 4)/9	0.80	6/8	1/8	75	this paper
12	{45,0},{40,20}	(5*k* - 4)/9	0.80	5/8	0/8	75	this paper
13	{45,0},{40,20}	(5*k* - 4)/9	0.80	6/8	5/8	75	this paper
14	{45,0},{40,20}	(5*k* - 4)/9	0.80	6/8	0/8	75	this paper
15	{45,0},{42,12}	(5*k* - 4)/15	0.80	4/8	3/8	75	this paper
16	{45,0},{42,12}	(5*k* - 4)/15	0.80	6/8	8/8*	75	this paper
17	{45,0},{42,12}	(5*k* - 4)/15	0.80	6/8	8/8*	75	this paper
18	{45,0},{42,12}	(5*k* - 4)/15	0.80	6/8	2/8	75	this paper
19	{100,20},{80,80}^{T}	(4*k* - 3)/5	0.75	2/10	0/10	9	Straub (1995)
20	{5,-1},{3,3}	(6*k* - 4)/5	0.67	5/12	3/12	10	Friedman (1996)
21	{5,-1},{4,1}	(3*k* - 2)/5	0.67	7/12	3/12	10	Friedman (1996)
22	{5,-1},{4,1}	(3*k* - 2)/5	0.67	7/12	3/12	16	Friedman (1996)
23	{5,-1},{4,1}	(3*k* - 2)/5	0.67	6/10	3/10	16	Friedman (1996)
24	{5,-1},{4,1}	(3*k* - 2)/5	0.67	2/12	3/12	10	Friedman (1996)
25	{5,-1},{4,1}	(3*k* - 2)/5	0.67	3/12	2/12	10	Friedman (1996)
26	{80,10},{70,30}^{T}	(3*k* - 2)/8	0.67	9/10	9/10	9	Straub (1995)
27	{100,20},{60,60}^{T}	(4*k* - 2)/5	0.50	7/10	10/10	9	Straub (1995)
28	{5,0},{4,1}	(2*k* - 1)/5	0.50	6/12	9/12	10	Friedman (1996)
29	{55,25},{35,35}^{T}	(6*k* - 2)/11	0.33	9/10	10/10	9	Staub (1995)

R(k) - scaled earnings difference given *k*, the probability of *X*: *R(k)* = ({*k* A + (1-*k*) C} - {*k* B + (1-*k*) D})/A
*k** - separatrix, zero earnings difference, mixed strategy equilibrium.
T - payoff dominant equilibrium in the lower right cell of subjects earnings table.
* - Separatrix crossings between initial and terminal outcome.

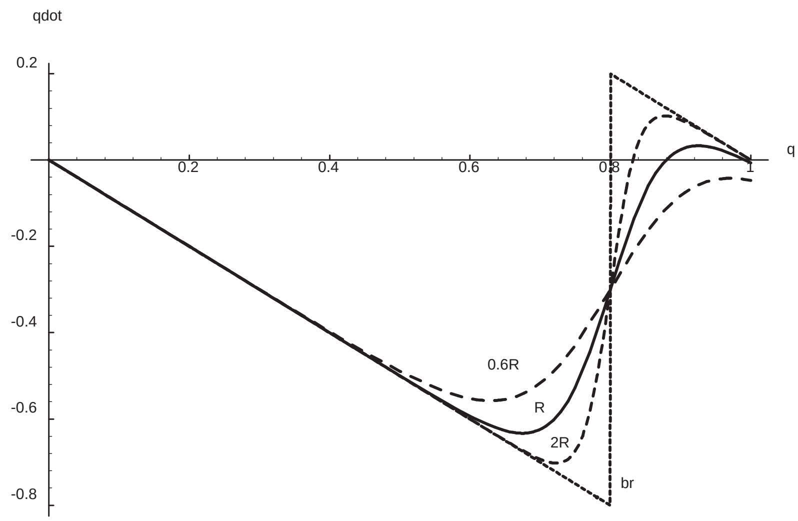


Figure 4: One population continuous time best response and logit response dynamics (λ=1).

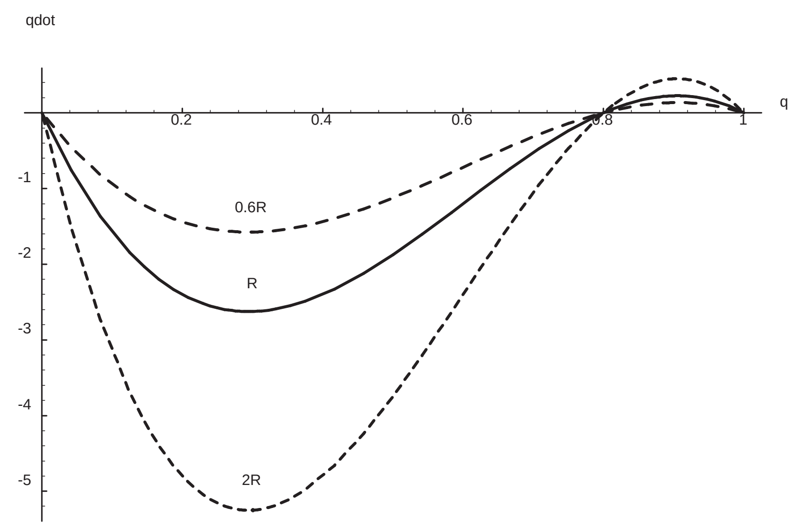
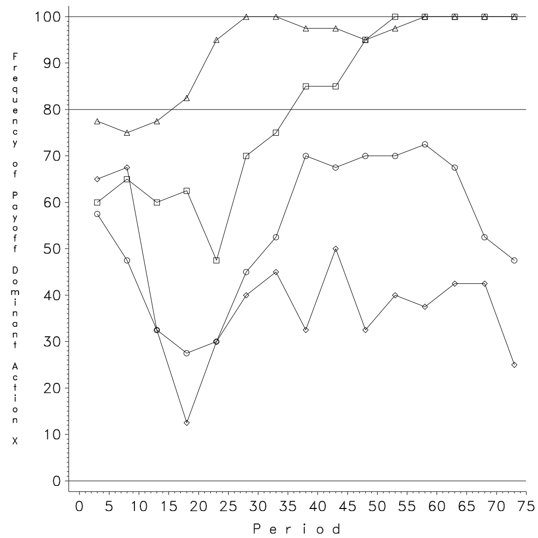
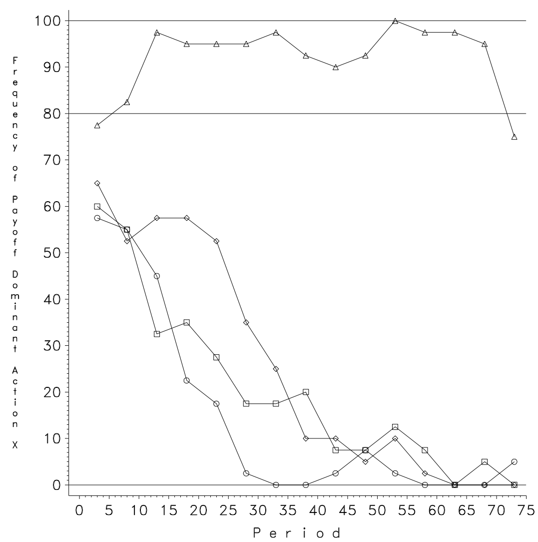


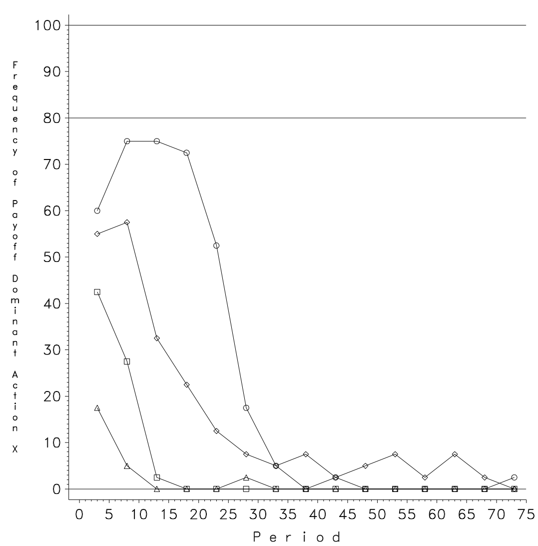
Figure 5: Continuous Time Replicator Dynamic for game *RD* (dashed line) and *PD* (solid line).



A. 0.6R Cohorts



B. R Cohorts



C. 2R Cohorts

Figure 6: five period means of *X* frequency.

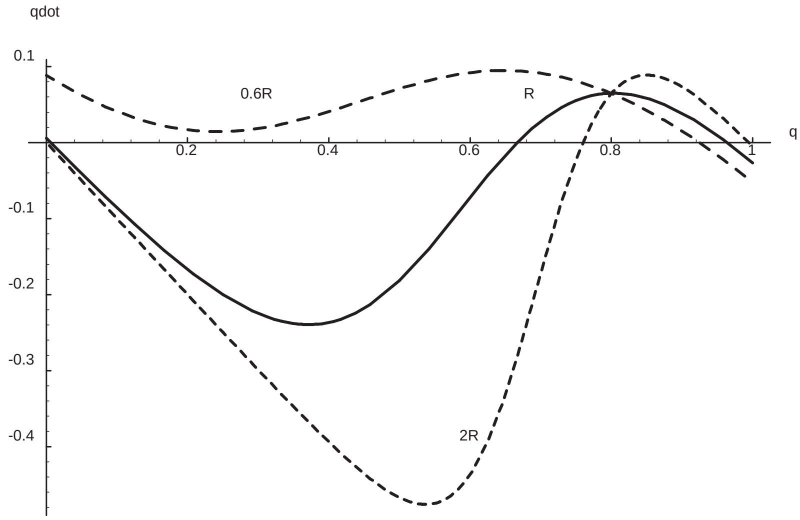


Figure 7: Logit response dynamic based on estimated EFP model.