Individual Differences in EWA Learning with Partial Payoff Information

Teck H. Ho
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6366

Colin F. Camerer
Division of Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91125

Xin Wang
Graduate School of Industrial Administration
Carnegie Mellon University
Pittsburgh, PA 15213

March 13, 2002

Individual Differences in EWA Learning with Partial Payoff Information

Abstract

We extend EWA learning to games in which only the set of possible foregone payoffs from unchosen strategies are known. We assume players estimate unknown foregone payoffs from a strategy, by substituting the last payoff actually received from that strategy, or by clairvoyantly guessing the actual foregone payoff. Either assumption improves predictive accuracy of EWA. Learning parameters are also estimated separately for each player. Players cluster into two separate subgroups, which do not correspond to traditional special cases of reinforcement or belief learning.

1 Introduction

How does an equilibrium emerge in a game? After largely avoiding this difficult question for many decades, we are now beginning to understand how equilibration occurs, empirically. There are two general approaches to understand learning: Population models and individual models.

Population models make predictions about how the aggregate behavior in a population will change as a result of aggregate experience. For example, in replicator dynamics, a population's propensity to play a certain strategy will depend on its 'fitness' (payoff) relative to the mixture of strategies played previously. Models like this submerge differences in individual learning paths.

Individual learning models allow each person to choose differently, depending on the experiences each person has. Our 'experience-weighted attraction' (EWA) model, for example, assumes that people learn by decaying experience-weighted lagged attractions, updating them according to received payoffs or weighted foregone payoffs, and normalizing. This very general approach includes the key features of reinforcement and belief learning (including Cournot and fictitious play), and predicts more accurately than those simple models, out-of-sample, in many different games (see Camerer, Ho, and Chong, in press for a comprehensive list).

In this paper, we the applicability of EWA in two ways.

First, we allow different players to have different learning parameters. In many previous empirical applications, players are assumed to have a common learning rule (exceptions include Cheung and Friedman, 1997; Stahl, 2000; Broseta, in press). Allowing heterogeneous parameter values is an important step for three reasons. (i) While it seems very likely that detectable heterogeneity exists, it is conceivable that allowing heterogeneity does not improve fit much, in which case 'representative agent' modeling with common parameter values is a useful approximation. (ii) If players are heterogeneous, it is likely that players fall into distinct clusters, perhaps corresponding to familiar learning rules like fictitious play or reinforcement learning, or to some other kinds of clusters not yet identified.[1] (iii)If players are heterogeneous, then it is possible that a single parameter estimated from a homogeneous representative- agent model

[1]Camerer and Ho (1998) allowed two separate configurations of parameters (or 'segments') to see whether the superior fit of EWA was due to its ability to mimic a population mixture of reinforcement and belief learners, but they found that this was clearly not so. The current study serves as another test of this possibility, with more reliable estimation of parameters for all players.

will misspecify the mean of the distribution of parameters across individuals. We test for any such bias by comparing the mean of individual estimates with the single representative-agent estimate.

Second, most theories of learning assume that players know the foregone payoffs to strategies they did not choose. Theories differ in the extent to which unchosen strategies are reinforced by foregone payoffs. (For example, belief learning theories are essentially just generalized reinforcement theories in which unchosen strategies are reinforced as strongly as chosen strategies are.) But then, as Vriend (1997) first noted, how does learning occur when players are not sure what foregone payoffs are? This is a crucial question for applying these theories to naturally-occurring situations in which the modeler may not know the foregone payoffs. In this paper we suggest two ways to add learning about unknown foregone payoffs ('payoff learning') to describe learning in low-information environments.

The basic results can be easily stated. We estimated individual-level EWA parameters for 60 subjects who played a normal-form centipede game (with extensive-form feedback) 100 times (see Nagel and Tang, 1998). Parameters do differ systematically across individuals. While parameter estimates do not cluster naturally around the values predicted by belief or reinforcement models, they do cluster in a striking way into learning in which attractions cumulate past payoffs, and learning in which attractions are averages of past payoffs.

Two payoff learning models are used to describe how subjects estimate foregone payoffs, then use these estimates to reinforce strategies whose foregone payoffs are not known precisely. Both are substantial improvements over the default assumption that these strategies are not reinforced at all (e.g., reinforcement models). The best model is the one in which subjects update unchosen strategies with perfect guesses of their foregone payoffs.

2 EWA Learning with Partial Payoff Information

2.1 The Basic EWA Model

Experience-weighted attraction learning was introduced to hybridize elements of reinforcement and belief-based approaches to learning and includes familiar variants of both as special cases. This section will highlight only the most important features of the model. Further details are available in Camerer and Ho (1999).

In EWA learning, strategies have attraction levels which are updated according to either the payoffs the strategies actually provided, or some fraction of the payoffs unchosen strategies would have provided. These attractions are decayed or depreciated each period, and also normalized by a factor which captures the (decayed) amount of experience players have accumulated. Attractions to strategies are then related to the probability of choosing those strategies using a response function which guarantees that more attractive strategies are played more often.

EWA was originally designed to study n-person normal form games. The players are indexed by i ($i = 1, 2, \ldots, n$), and each one has a strategy space $S_i = \{s_i^1, s_i^2, \ldots, s_i^{m_i-1}, s_i^{m_i}\}$, where s_i denotes a pure strategy of player i. The strategy space for the game is the Cartesian products of the S_i, $S = S_1 \times S_2 \times \ldots \times S_n$. Let $s = (s_1, s_2, \ldots, s_n)$ denote a strategy combination consisting of n strategies, one for each player. Let $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$ denote the strategies of everyone but player i. The game description is completed with specification of a payoff function $\pi_i(s_i, s_{-i}) \in \Re$, which is the payoff i receives for playing s_i when everyone else is playing the strategy specified in the strategy combination s_{-i}. Finally, let $s_i(t)$ denote i's actual strategy choice in period t, and $s_{-i}(t)$ the vector chosen by all other players. Thus, player i's payoff in period t is given by $\pi_i(s_i(t), s_{-i}(t))$.

2.2 Updating Rules

The EWA model updates two variables after each round. The first variable is the experience weight $N(t)$, which is like a count of 'observation-equivalents' of past experience and is used to weight lagged attractions when they are updated. The second variable is $A_i^j(t)$, the attraction of a strategy after period t has taken place. The variables $N(t)$ and $A_i^j(t)$ begin with initial values $N(0)$ and $A_i^j(0)$. These prior values can be thought of as reflecting pregame experience, either due to learning transferred from different games or due to introspection.

Updating after a period of play is governed by two rules. First, experience weights are updated according to

$$N(t) = \phi \cdot (1 - \kappa) \cdot N(t-1) + 1, \quad t \geq 1. \tag{2.1}$$

The second rule updates the level of attraction. A key component of the updating is the payoff that a strategy either yielded, or would have yielded, in a period. The model weights hypothetical payoffs that unchosen strategies would have earned by a parameter δ, and weights payoff actually received, from chosen strategy $s_i(t)$, by an additional $1-\delta$ (so it receives a total weight of 1). Using an indicator function $I(x,y)$ which equals 1 if $x=y$ and 0 if $x \neq y$, the weighted payoff for i's j^{th} strategy can be written $[\delta+(1-\delta)\cdot I(s_i^j, s_i(t))]\cdot \pi_i(s_i^j, s_{-i}(t))$. The rule for updating attraction sets $A_i^j(t)$ to be a depreciated, experience-weighted lagged attraction, plus an increment for the received or foregone payoff, normalized by the new experience weight. That is,

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + [\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))}{N(t)}. \tag{2.2}$$

The factor ϕ is a discount factor that depreciates previous attractions. The parameter κ adjusts whether the experience weight depreciates more rapidly than the attractions. Notice that the steady-state value of $N(t)$ is $\frac{1}{1-\phi(1-\kappa)}$ (and does not depend on $N(0)$). In the estimation we usually impose the restriction $N(0) \leq \frac{1}{1-\phi(1-\kappa)}$ which guarantees that the experience weight rises over time, so the relative weight on new payoffs falls and learning slows down.

Finally, attractions must be related to the probabilities of choosing strategies in some way. Obviously we would like $P_i^j(t)$ to be monotonically increasing in $A_i^j(t)$ and decreasing in $A_i^k(t)$ (where $k \neq j$). Three forms have been used in previous research: A logit or exponential form, a power form, and a normal (probit) form. The various probability functions each have advantages and disadvantages. We prefer the logit form

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}} \tag{2.3}$$

because it allows negative attractions and fits a little better in a direct comparison with the power form (Camerer and Ho, 1998). The parameter λ measures sensitivity of players to differences among attractions. When λ is small, probabilities are not very sensitive to differences in attractions (when $\lambda = 0$ all strategies are equally likely to be chosen). As λ in creases, it converges to a best-response function in which the strategy with the highest attraction is always chosen.

2.3 The Cumulative Reinforcement Special Case of EWA

One special case of EWA is choice reinforcement models in which strategies have levels of reinforcement or propensity which are depreciated and incremented by received payoffs. In the model of Harley (1981) and Roth and Erev (1995), Erev and Roth (1998), for example

$$R_i^j(t) = \begin{cases} \phi \cdot R_i^j(t-1) + \pi_i(s_i^j, s_{-i}(t)) & \text{if } s_i^j = s_i(t), \\ \phi \cdot R_i^j(t-1) & \text{if } s_i^j \neq s_i(t). \end{cases} \tag{2.4}$$

Using the indicator function, the two equations can be reduced to one:

$$R_i^j(t) = \phi \cdot R_i^j(t-1) + I(s_i^j, s_i(t)) \cdot \pi_i(s_i^j, s_{-i}(t)). \tag{2.5}$$

It is easy to see that this updating formula is a special case of the EWA rule, when $\delta = 0$, $N(0) = 1$, and $\kappa = 1$. The reinforcement model is extremely simple; in fact, it is too simple and has been largely abandoned by cognitive psychologists studying humans in favor of algorithmic information processing models or connectionist neural networks. Nonetheless, the adequacy of the model can be tested empirically by setting the parameters to their restricted values and seeing how much fit is compromised (adjusting, of course, for degrees of freedom).

2.4 The Averaged Reinforcement Special Case of EWA

In another kind of reinforcement, attractions are averages of previous attractions, and reinforcements, rather than cumulations (e.g. Sarin and Vahid, 1997; Mookerjhee and Sopher, 1994, 1997). For example

$$R_i^j(t) = \phi \cdot R_i^j(t-1) + (1-\phi) \cdot I(s_i^j, s_i(t)) \cdot \pi_i(s_i^j, s_{-i}(t)). \tag{2.6}$$

A little algebra shows that this updating formula is a special case of the EWA rule, when $\delta = 0$, $N(0) = \frac{1}{1-\phi}$, and $\kappa = 0$.

2.5 The Belief-Based Special Case of EWA

In belief-based models, adaptive players base their responses on beliefs formed by observing their opponents' past plays. While there are many ways of forming beliefs, we consider a fairly general 'weighted fictitious play' model, which includes fictitious play (Brown, 1951; Fudenberg and Levine, 1998) and Cournot best-response (Cournot, 1960) as special cases.

In weighted fictitious play, players begin with prior beliefs about what the other players will do, which are expressed as ratios of counts to the total experience. Denote total experience by $N(t) = \sum_{k=1}^{m_{-i}} N_{-i}^{k}(t)$.[2] Express the probability that others will play strategy k as $B_{-i}^{k}(t) = \frac{N_{-i}^{k}(t)}{N(t)}$, with $N_{-i}^{k}(t) \geq 0$ and $N(t) > 0$.

Beliefs are updated by depreciating the previous counts by ϕ, and adding one for the strategy combination actually chosen by the other players. That is,

$$B_{-i}^{k}(t) = \frac{\phi \cdot N_{-i}^{k}(t-1) + I(s_{-i}^{k}, s_{-i}(t))}{\sum_{h=1}^{m_{-i}}[\phi \cdot N_{-i}^{h}(t-1) + I(s_{-i}^{h}, s_{-i}(t))]}. \tag{2.7}$$

This form of belief updating weights the belief from one period ago ϕ times as much as the most recent observation, so ϕ can be interpreted as how quickly previous experience is discarded.[3] When $\phi = 0$ players weight only the most recent observation (Cournot dynamics); when $\phi = 1$ all previous observations count equally (fictitious play).

Given these beliefs, we can compute expected payoffs in each period t,

$$E_{i}^{j}(t) = \sum_{k=1}^{m_{-i}} B_{-i}^{k}(t)\pi(s_{i}^{j}, s_{-i}^{k}). \tag{2.8}$$

The crucial step is to express period t expected payoffs as a function of period $t-1$ expected payoffs. This yields:

$$E_{i}^{j}(t) = \frac{\phi \cdot N(t-1) \cdot E_{i}^{j}(t-1) + \pi(s_{i}^{j}, s_{-i}(t))}{\phi \cdot N(t-1) + 1}. \tag{2.9}$$

Expressing expected payoffs as a function of lagged expected payoffs, the belief terms disappear into thin air. This is because the beliefs are only used to compute expected payoffs, and when beliefs are formed according to weighted fictitious play, the expected payoffs which result can

[2]Note that $N(t)$ is not subscripted because the count of frequencies is assumed, in our estimation, to be the same for all players. Obviously this restriction can be relaxed in future research.

[3]Some people interpret this parameter as an index of 'forgetting', but this interpretation is misleading because people may recall the previous experience perfectly (or have it available in 'external memory' on computer software) but they will deliberately discount old experience if they think new information is more useful in forecasting what others will do.

also be generated by generalized reinforcement according to previous payoffs. More precisely, if the initial attractions in the EWA model are expected payoffs given some initial beliefs (i.e., $A_i^j(0) = E_i^j(0)$), $\kappa = 0$, and foregone payoffs are weighted as strongly as received payoffs ($\delta = 1$), then EWA attractions are exactly the same as expected payoffs.

This demonstrates a close kinship between reinforcement and belief approaches. Belief learning is nothing more than generalized attraction learning in which strategies are reinforced equally strongly by actual payoffs and foregone payoffs, attractions are weighted averages of past attractions and reinforcements, and initial attractions spring from prior beliefs. This relation is quite surprising and previous scholars did not notice it. For example, Selten (in press) wrote,"In rote [reinforcement] learning success and failure directly influence the choice probabilities. Belief learning is very different. Here experiences strengthen or weaken beliefs. Belief learning has only an indirect influence on behavior. (emphasis ours). The EWA model shows that when beliefs are formed by weighted fictitious play, the "indirect influence" Selten referred to is exactly the same as a certain kind of direct influence.

2.6 Interpreting EWA

The EWA parameters have the following psychological interpretations.

1. The parameter δ measures the relative weight given to foregone payoffs, compared to actual payoffs, in updating attractions. It can be interpreted as a kind of "imagination" of foregone payoffs, or responsiveness to foregone payoffs (when δ is larger players move more strongly toward ex post best responses). We call it "consideration" of foregone payoffs.

2. The parameter ϕ is naturally interpreted as depreciation of past attractions, $A(t)$. In a game-theoretic context, ϕ will be affected by the degree to which players realize other players are adapting, so that old observations on what others did become less and less useful. Then ϕ can be interpreted as an index of (perceived) change.

3. The parameter κ determines the growth rate of attractions, which in turn affects how sharply players converge. When $\kappa = 1$ then $N(t) = 1$ (for $t > 0$) and the denominator in the attraction updating equation disappears. Thus, attractions cumulate past payoffs as

quickly as possible. When $\kappa = 0$, attractions are weighted averages of lagged attractions and past payoffs, where the weights are $\phi \cdot N(0)$ and 1.

In the logit model, whether attractions cumulate payoffs, or average them, is important because only the difference among the attractions matters for their their relative probabilities of being chosen. (A constant added to all the attractions divides out in the logit form, so only the difference between two strategy's attractions affects their relative probability.) If attractions can grow and grow, as they can when $\kappa = 1$ then the differences in strategy attractions can be very large. This implies that, for a fixed response sensitivity, λ, the probabilities can be spread farther apart; convergence to playing a single strategy almost all the time can be sharper. If attractions cannot grow outside of the payoff bounds, when $\kappa = 0$ then convergence cannot produce choice probabilities which are so extreme. Thus, we think of κ as an index of the degree of commitment to one choice or another (it could also be thought of as a convergence index, or confidence). When κ is large then attractions are cumulating payoffs rapidly, and the probability of choosing one strategy will grow closer to one. When κ is small, the attractions are not growing as rapidly and probability will be spread more evenly across strategies.

4. The term $A_i^j(0)$ represents the initial attraction, which might be derived from some analysis of the game, from selection principles or decision rules, from surface similarity between strategies in the game being played and strategies which were successful in similar games, etc. Belief models impose strong restrictions on $A_i^j(0)$ by requiring initial attractions to be derived from prior beliefs.[4] Additionally, they require attraction updating with $\delta = 1$ and $\kappa = 0$. EWA allows one to separate these two processes: Players could have arbitrary initial attractions but begin to update attractions in a belief-learning way after they gain experience.

5. The initial-attraction weight $N(0)$ is in the EWA model to allow players in belief-based models to have an initial prior which has a strength (measured in units of actual experience). In EWA, $N(0)$ is therefore naturally interpreted as the strength of initial attractions, relative to incremental changes in attractions due to actual experience and

[4]This requires, for example, that weakly dominated strategies will always have (weakly) lower initial attractions than dominant strategies. EWA allows more flexibility. For example, players might choose randomly at first, choose what they chose previously in a different game, or set a strategy's initial attraction equal to its minimum payoff (the minimax rule) or maximum payoff (the maximax rule). All these decision rules generate initial attractions which are not generally allowed by belief models, but are permitted in EWA because $A_i^j(0)$ are flexible.

payoffs. The effect of $N(0)$ is easiest to see by fixing $\delta = 1$ for simplicity and directly computing the attraction after two periods, $A_i^j(2)$, which gives

$$A_i^j(2) = \frac{\phi^2 \cdot A_i^j(0) \cdot N(0) + \phi \cdot \pi_i(s_i^j, s_{-i}(1)) + \pi_i(s_i^j, s_{-i}(2))}{\phi^2 \cdot (1-\kappa)^2 \cdot N(0) + \phi \cdot (1-\kappa) + 1}. \tag{2.10}$$

The parameter ϕ captures the declining weight placed on payoffs from more distant periods of actual experience (that is, the period 1 payoff $\pi_i(s_i^j, s_{-i}(1))$ is weighted ϕ but the period 2 payoff $\pi_i(s_i^j, s_{-i}(2))$ is not). Like previous payoffs, the initial attraction is also weighted by a power of ϕ (ϕ^2, because it 'happened' two periods earlier), but is also weighted by $N(0)$. Thus, the parameter $N(0)$ captures the special weight placed on the initial attractions, compared to increments in attraction due to actual (or foregone) payoffs. If $N(0)$ is small then the effect of the initial attractions wears off very quickly (compared to the effect of actual experience). If $N(0)$ is large then the effect of the initial attractions persists.[5]

In previous research, the EWA model has been estimated on several samples of experimental data, and estimates have been used to predict out-of-sample. Compared to the belief and reinforcement special cases, EWA fits better in weak- link coordination games (Camerer and Ho, 1998; out-of-sample accuracy was not measured) and predicts better out of sample in median-action coordination games and dominance solvable "p-beauty contests" (Camerer and Ho, 1999), call markets (Hsia, 1998),"unprofitable games" (Morgan and Sefton, forthcoming), partially-dominance-solvable R&D games (Rapoport and Almadoss, 2000), and in unpublished estimates we made in other "continental divide" coordination games (Van Huyck, Cook, and Battalio, 1997). EWA only predicted substantially worse than belief learning in some constant-sum games (Camerer and Ho, 1999), and has never predicted substantially worse than reinforcement learning.

To help illustrate how EWA hybridizes features of other theories, Figure 1 shows a three-dimensional parameter space– a cube– in which the parameters are δ, ϕ, and κ. Traditional belief and reinforcement theories assume that learning parameters are located on specific edges of the cube. For example, cumulative reinforcement theories require low consideration ($\delta = 0$) and high commitment ($\kappa = 1$). (Note that the combination of low consideration and high commitment may be the worst possible combination, since such players can get quickly locked

[5]This enables one to test equilibrium theories as a special kind of (non)-learning theory with $N(0)$ very large and initial attractions equal to equilibrium payoffs.

in to strategies which are far from best responses.) Belief models are represented by points on the edge where consideration is high ($\delta = 1$) but commitment is low ($\kappa = 0$). This constrains the ability of belief models to produce sharp convergence, in coordination games for example (Camerer and Ho, 1998,1999). Cournot best- response and fictitious play learning are vertices at the ends of the belief- model edge.

Looking at Figure 1, it is unclear why players' parameter values might cluster on those edges or vertices (as opposed to other areas, or the interior of the cube). In fact, we shall see below that there is no prominent clustering in the regions corresponding to familiar belief and reinforcement models, but there is substantial clustering near the faces where commitment is either low ($\kappa = 0$) or high ($\kappa = 1$).

2.7 EWA Extensions to Partial Payoff Information

Our approach to explaining learning in environments with partial payoff information is to assume that players form some guess about what the foregone payoff might be, then plug it into the attraction updating equation. This adds no free parameters to the model.

First define the estimate of the foregone payoff as $\hat{\pi}_i(s_i^j, t)$ (and $\hat{\pi}$ is just the known foregone payoff when it is known). Note that $\hat{\pi}_i(s_i^j, t)$ does <u>not</u> generally depend on $s_{-i}(t)$ because, by definition, if the other players' strategy was observed then the foregone payoff would be known. Hence, when the foregone payoff is not known, updating is done according to

$$N_i^j(t) = \rho \cdot N_i^j(t-1) + 1, \quad t \geq 1. \tag{2.11}$$

and

$$A_i^j(t) = \frac{\phi \cdot N_i^j(t-1) \cdot A_i^j(t-1) + \delta \cdot \hat{\pi}_i(s_i^j, t)}{N_i^j(t)}. \tag{2.12}$$

We try two separate specifications of $\hat{\pi}(s_i^j, t)$: Last actual payoff updating, and payoff clairvoyance. When players update according to the last actual payoff, they recall the last payoff they actually received from a strategy, and use that as an estimate of the foregone payoff. Formally,

$$\hat{\pi}_i(s_i^j, t) = \begin{cases} \pi_i(s_i^j, s_{-i}(t)) & \text{if } s_i^j = s_i(t), \\ \hat{\pi}_i(s_i^j, t-1) & \text{otherwise.} \end{cases} \tag{2.13}$$

To complete the specification, the estimates $\hat{\pi}_i(s_i^j, 0)$ are initialized as the average of all the possible elements of the set of foregone payoffs. For example, suppose player A chooses 7 and player B chooses 8 or higher. Since player A "took first" she receives a payoff of 32, and she knows that if she chose 9 instead, she would receive either 11, if player B chose 8, or 64 if player B chose 10, 12, or 14. In this case we would initialize $\hat{\pi}_i(9, 0) = (11 + 64)/2$. Notice that we average only the unique elements of the payoff set, not each payoff associated with every strategy pair. That is, even though 64 would result if player A chose 8 and B chose 10, 12, or 14, we only use the payoff 64 once, not three times, in computing the initial $\hat{\pi}$.

Updating using the last actual payoff is cognitively economical because it requires players to remember only the last payoff they received. Furthermore, it enables them to adjust rapidly when other players' behavior is changing, by immediately discounting all previous received payoffs and focussing only the most recent one.

On the other hand, if one thinks of the last actual payoff as an implicit forecast of what payoff is likely to have been the 'true' foregone one, then it may be a poor forecast when the last actual payoff was received many periods ago, or if subjects have hunches about which foregone payoff they would have gotten which are more accurate than distant history. Therefore, we consider an opposite assumption as well– 'payoff clairvoyance'. Under payoff clairvoyance, $\hat{\pi}_i(s_i^j, t) = \pi_i(s_i^j, s_{-i}(t))$. That is, players accurately guess exactly what the foregone payoff would have been even though, strictly speaking, there is no way they could know this.

There are other assumptions one could consider, but these two represent opposite points in terms of information use. The last-actual-payoff scheme recalls only observed history and does not try to improve upon it (as a forecast); consequently, it can also be applied when players do not even know the set of possible foregone payoffs. The payoff-clairvoyance scheme uses knowledge which the subject is not told (but could conceivably figure out). We report estimates and fit measures for both models.

3 Data

Nagel and Tang (1998) (NT) studied learning in an extensive-form centipede game. In their game, an Odd player has the opportunity to take the majority of a growing 'pie' at odd numbered decision nodes {1,3,5,7,9,11,13}; the Even player has the opportunity to take at nodes {2,4,6,8,10,12,14}. Each player chooses when to take by choosing a number. The lower of the two numbers determines when the pie stops growing and how much each player gets. The person who chooses the lower number always gets more. Table 1 shows the payoffs to the players from taking at each node. (Points are worth .005 deutschemarks).

They conducted five sessions with 12 subjects in each, playing 100 rounds in a random-matching fixed-role protocol. A crucial design feature is that while the players choose normal-form strategies, they are given extensive-form feedback. That is, each pair of subjects is only told the lower number chosen in each round, corresponding to the time at which the pie is taken and the game stops. The player choosing the lower number does not know the higher number. For example, if Odd chooses 5, takes first, and earns 16, she is not sure whether she would have earned 6 by taking later, at node 7 (if Even's number was 6) or whether she would have earned 32 (if Even had taken at 8 or higher), because she only knows that Even's choice was higher than 5. This ambiguity about foregone payoffs is an important challenge for implementing learning models.

Table 2 shows the overall frequencies of choices (pooled across the five sessions, which are similar). Most players choose numbers from 7 to 11.

If a subject's number was the lower one (i.e., they chose "take"), there is a strong tendency to choose the same number, or a higher number, on the next round. This can be seen in the transition matrix Table 3, which shows the relative frequency of choices in round t+1 as a function of the choice in round t, for players who 'take' in round t (choosing the lower number). For example, the top row shows that when players choose 2 and take, they choose 2 in the next round 28% of the time, but 8% choose 4 and 32% choice 6, which is the median choice (and is underlined). For choices below 6, the median choice in the next period is always higher. The overall tendency for players who chose "take" to choose numbers which increase, decrease, or are unchanged are shown in Figure 2a. Note that most "takers" then choose numbers which increase, but this tendency shrinks over time.

Table 4 shows the opposite pattern for players who choose the higher number and "pass"–

they tend to choose lower numbers. In addition, as the experiment progressed this pattern of transitions became weaker (more subjects do not change at all), as Figure 2a shows.

NT consider several models. Four are benchmarks which assume no learning: Nash equilibrium (players pick 1 and 2), quantal response equilibrium (McKelvey and Palfrey, 1995) random play, and an individual observed-frequency model which uses each player's observed frequencies of choices over all 100 rounds. NT test choice-reinforcement of the Harley-Roth-Erev RPS type and implement weighted fictitious play in an odd way, which assumes players have population history information which they don't actually have. The equilibrium and weighted fictitious play predictions are terrible. This is not surprising because both theories predict either low numbers at the start, or steady movement toward lower numbers over time, which is obviously not present in the data. QRE and random guessing don't predict too badly, but the individual-frequency benchmark is the best of all. The RPS (reinforcement) models do almost as well as the best benchmark.

4 Estimation Methodology

The method of maximum likelihood was used to estimate the various model

parameters. We used the first 70% of the data to calibrate the models and the last 30% of the data to predict out-of-sample. Note well that out-of-sample forecasting completely removes any advantage more complicated models have over simpler ones which are special cases.

We first estimated a homogeneous single-representative agent model for reinforcement, belief, and four variants of EWA payoff learning. We then estimated the EWA models at the individual level for all 60 subjects. In the centipede game, each subject has seven strategies, numbers $1, 3, \ldots, 13$ for Odd subjects and $2, 4, \ldots, 14$ for even subjects. Since the game is asymmetric, the models for Odd and Even players were estimated separately. The log of the likelihood function for the single-representative agent EWA model is

$$LL(\delta, \phi, \kappa, \lambda, N(0), \tau) = \sum_{i=1}^{30} \sum_{t=2}^{70} log(P_i^{S_i(t)}(t)) \tag{4.1}$$

and for the individual level model for player i is:

$$LL(\delta, \phi, \kappa, \lambda, N(0), \tau) = \sum_{t=2}^{70} log(P_i^{S_i(t)}(t)) \tag{4.2}$$

where the probabilities $P_i^{S_i(t)}(t)$ are given by equation (2.3).

There is one substantial change from methods we previously used in Camerer and Ho (1999). We estimated initial attractions (common to all players) from the first period of actual data, rather than allowing them to be free parameters which are estimated as part of the overall maximization of likelihood.[6] We switched to this method because estimating initial attractions for each of the the large number of strategies chewed up too many degrees of freedom.

To search for regularity in the distributions of individual-level parameter estimates, we conducted a cluster analysis on the three most important parameters, δ, ϕ, and κ. We specified a number of clusters and searched iteratively for cluster means in the three-dimensional parameter space which maximizes the ratio of the distance between the cluster means and the average within-cluster deviation from the mean. We report results from two- cluster specifications, since they have special relevance for evaluating whether parameters cluster around the predictions of belief and reinforcement theories. Searching for a third cluster generally improved the fit very little.[8]

[6]Others have used this method too, e.g., Roth and Erev (1995). Formally, define the first-period frequency of strategy j in the population as f^j. Then initial attractions are recovered from the equations

$$\frac{e^{\lambda \cdot A^j(0)}}{\sum_k e^{\lambda \cdot A^k(0)}} = f^j, j = 1, \ldots, m \tag{4.3}$$

(This is equivalent to choosing initial attractions to maximize the likelihood of the first-period data, separately from the rest of the data, for a value of λ derived from the overall likelihood-maximization.) Some algebra shows that the initial attractions can be solved for, as a function of λ, by

$$A^j(0) - \frac{1}{m}\sum_j A^j(0) = \frac{1}{\lambda} ln(\tilde{f}^j), j = 1, \ldots, m \tag{4.4}$$

where $\tilde{f}^j = \frac{f^j}{(\Pi_k f^k)^{\frac{1}{m}}}$ is a measure of relative frequency of strategy j. We fix the strategy j with the lowest frequency to have $A^j(0) = 0$ (which is necessary for identification) and solve for the other attractions as a function of λ and the frequencies $\tilde{f}^j$.

Estimation of the belief-based model (a special case of EWA) is a little trickier. Attractions are equal to expected payoffs given initial beliefs; therefore, we searched for initial beliefs which optimized the likelihood of observing the first-period data.[7]

[8]Specifically, a three-segment model always leads to a tiny segment that contains either 1 or 2 subjects.

5 Results

We discuss the results in three parts: Basic estimation and model fits; comparison of two payoff-learning extensions, and individual-level estimates and uncovered clusters.

5.1 Basic estimation and model fits

Table 5 reports the log-likelihood of the various models, both in-sample and out-of-sample. The belief-based model is clearly worst by all measures. This is no surprise because the centipede game is dominance-solvable. Any belief learning should move players in the direction of lower numbers, but the numbers they choose rise slightly over time. The EWA-Payoff Clairvoyance is better than the EWA - Recent Actual Payoff. Reinforcement is worse than any of the EWA variants, and by about 50 points of log-likelihood out-of-sample when compared to the best EWA model. (It can also be strongly rejected in- sample using standard χ^2 tests.)

Another way to judge model fit is to see how well the EWA model estimates capture the basic patterns in the data. There are two basic patterns: (i) players who choose the lower number (and 'take earlier', in centipede jargon) tend to increase their number more often than they decrease it, and this tendency decreases over time; and (ii) players who choose the higher number ('taking later'), tend to decrease their numbers.

Figures 2a show these patterns in the data and Figures 2b-c show how well the EWA model describes and predicts these patterns. The EWA predictions are generally quite accurate. Note that if EWA were overfitting in the first 70 periods, accuracy would degrade badly in the last 30 periods (when parameter estimates are fixed and out-of-sample prediction begins); but it generally doesn't.

5.2 Payoff learning models

Tables 5 show measures of fit and parameter estimates for two different payoff learning models. The two models either use the last actual payoff or the clairvoyantly known payoff. Both payoff learning models perform better than reinforcement (which implicitly assumes that the estimated foregone payoff is zero, or gives it zero weight by setting $\delta = 0$). This illustrates that EWA can improve statistically on its special cases, even when foregone payoffs are not known.

By simply adding a payoff-learning assumption to EWA, the extended model outpredicts the simpler models. Building on our idea, the same superiority of EWA supplemented by payoff learning is shown by Anderson (1998) in bandit problems and Chen (1999), in a study of joint cost allocation.

The two payoff learning assumptions embody low and high degrees of player knowledge. The assumption that players recall only the last actual payoff– which may have been received many periods ago– means they ignore deeper intuitions about which of the possible payoffs might be the correct foregone one in the very last period. Oppositely, the payoff clairvoyance assumption assumes the players somehow figure out exactly which foregone payoff they would have gotten. Surprisingly, the payoff clairvoyance assumption predicts better. The right interpretation is surely not that subjects are truly clairvoyant, always guessing the true foregone payoff perfectly, but simply that their implicit foregone payoff estimate is sometimes closer to the truth than the last actual payoff is. For example, consider a player B who chooses 6 and has the lower of the two numbers. If she had chosen strategy 8 instead, she doesn't know whether the foregone payoff would have been 8 (if the other A subject chose 7), or 45 (if the A subject chose 9, 11, or 13). The payoff clairvoyance assumption says she knows precisely whether it would have been 8 or 45 (i.e., whether subject A chose 7, or chose 9 or more). While this requires knowledge she doesn't have, it only has to be a better guess than the last actual payoff she got from choosing 8.

5.3 Individual differences

The fact that Nagel and Tang's game lasted 100 trials enabled us to estimate individual-level parameters with some reliability (while imposing common initial attractions). Figures 3a-b show scatter plots of the 30 estimates from the EWA-Payoff Clairvoyance model in a three-parameter $\delta - \phi - \kappa$ space. Each point represents a triple of estimates for a specific player; a vertical projection to the bottom face of the cube helps the eye locate the point in space and measure its $\phi - \kappa$ values. Figure 3a shows Odd players and Figure 3b shows Even players. As shown, both figures are very similar and suggest that there is no difference between odd and even players. In addition, these parameter estimates are clearly grouped into two clusters.

Individuals do not particularly fall into clusters corresponding to any of the familiar special cases (compare Figure 1 and Figures 3a-b). For example, only two out of sixty subjects are near the cumulative reinforcement line $\delta = 0, \kappa = 1$ (the "bottom back wall"). Similarly, eight out of sixty subjects are clustered near the weighted fictitious play line where $\delta = 1, \kappa = 0$.

Table 6 shows the mean of the parameter estimates, along with standard deviations across subjects, for the EWA model. Results for Odd and Even players are reported separately, because the game is not symmetric. The separate reporting also serves as a kind of robustness check, since there is no reason to expect their learning parameters to be systematically different; and in fact, the parameters are quite similar for the two groups of subjects. The consideration parameter δ ranges from .44 to .47, the change parameter ϕ varies only a little, from .90 to .91, and the commitment parameter κ from .34 to 0.38. The standard deviations of δ and κ are large, which indicates the presence of substantial heterogeneity.

We normalize the three parameter estimates by substracting them from their means and dividing them by their ranges. (Milligan and Cooper (1985) studied eight normalization methods and showed that this was the best). We use a k-mean clustering algorithm and apply Calinski and Harabasz (1974)'s method to determine the the optimal number of clusters.[9] For both players , the optimal number of clusters is two. The means and within-cluster standard deviations of parameter values are given in Table 7. The subjects can be sorted into two clusters, one is twice the size of the other. Both clusters tend to have δ around .45 and ϕ around .8-.9; however, in one cluster κ is very close to zero and in the other cluster κ is close to one. Graphically, subjects tend to cluster on the front wall representing low ($\kappa = 0$) commitment, and the back wall representing high ($\kappa = 1$) commitment.

In earlier work we typically estimated a model in which all players are assumed to have the same learning parameters (i.e., a representative agent approach). Econometrically, it is possible that a parameter estimated with that approach will give a biased estimate of the population mean of the same parameter across individuals, when there is heterogeneity. We can test for this danger directly by comparing the mean of parameter estimates in Table 6 with estimates from a single-agent analysis assuming homogeneity. The estimates are not wildly dissimilar, but there

[9]The method picks the optimal number of cluster by maximizing the following index:

$$\frac{trace(B)/g - 1}{trace(W)/n - g} \tag{5.1}$$

where $n = 30$ is the number of odd (or even) subjects, g is the number of clusters, B is the between sum of squares cross product matrice, W is the within sum of squares cross product matrice, and the trace function represents the sum of all diagnoal elements in a matrice. Milligan and Cooper (1985) reported that this procedure generally performs the best and recovers the true clusters in more than 90% of the cases.

are some slight biases which are worth noting. Specifically, δ tends to be under-estimated by the representative-agent model, and is estimated to be 0.28 when the mean of sample estimates is around .45. Furthermore, the parameter κ from the single-agent model tends to take the extreme value of 0 or 1, when the sample means are around .36. Since there is substantial heterogeneity among subjects– the clusters show that subjects tend to have high κ's near 1, or low values near 0– the single-agent model seems to use a kind of 'majority rule' in loglikelihood and chooses one extreme value or the other, rather than choosing the sample mean.

6 Conclusions

In this paper, we extend our experience-weighted attraction (EWA) learning model to games in which players know the set of possible foregone payoffs from unchosen strategies, but do not precisely which payoff they would have gotten. This extension is crucial for applying the model to naturally-occurring situations in which the modeller (and even the players) do not know much about the foregone payoffs.

To model how players respond to unknown foregone payoffs, we allowed players to learn about them by substituting the last payoffs received when those strategies were actually played, or by clairvoyantly guessing the actual foregone payoffs. Our results show that these payoff-learning EWA models fit and predict better than reinforcement and belief learning. The clairvoyant-guessing model fits slightly better than the last-actual-payoff model.

We also estimated parameters separately for each individual player. The individual estimates showed there is substantial heterogeneity, but individuals could not be sharply clustered into either reinforcement or belief- based models (though several did have weighted fictitious play learning parameters). They could, however, be clustered into two distinct subgroups, corresponding to averaging and cumulating of attraction. Compared to the means of individual level estimates, the parameter estimates from the single-agent model have a tendency to modestly underestimate δ and take extreme values for κ when the average is around 0.36.

Future research should apply these payoff-learning specifications, and others, to environments in which foregone payoffs are unknown (see Anderson, 1998, and Chen, 1999). If we can find a payoff-learning specification which fits reasonably well across different games, then EWA with payoff learning can be used on naturally-occurring data sets, taking the study of learning outside the lab and providing new challenges.

References

[1] Anderson, C. 1998. Learning in bandit problems. Caltech Working Paper.

[2] Brown, G. 1951. Iterative Solution of Games by Fictitious Play. In *Activity Analysis of Production and Allocation*, New York: John Wiley & Sons.

[3] Broseta, B. In press. Adaptive learning and equilibrium selection in experimental coordination games: An ARCH(1) approach. *Games and Economic Behavior*.

[4] Calinski, T., J. Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics*, 3, 1-27.

[5] Camerer, C., T-H. Ho. 1998. Experience-weighted learning in coordination games: probability rules, heterogeneity, and time variation. *Journal of Mathematical Psychology*, 42, 305-326.

[6] Camerer, C., T-H. Ho. 1999. Experience-Weighted attraction learning in normal-form games. *Econometrica*, 67, 827-873.

[7] Camerer, C., T-H. Ho, and J-K. Chong. In press. Sophisticated EWA learning and strategic teaching in repeated games. *Journal of Economic Theory*.

[8] Chen, Y. 1999. Joint cost allocation in asynchronously updated systems. Department of Economics Working Paper, University of Michigan.

[9] Cheung, Y-W., D. Friedman. 1997. Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, 19, 46-76.

[10] Cournot, A. 1960 *Recherches sur les principes mathematiques de la theorie des richesses.* Translated into English by N. Bacon as Researches in the Mathematical Principles of the Theory of Wealth. London: Haffner.

[11] Erev, I., A. Roth. 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed-strategy equilibria. *American Economic Review*, 88, 848-881.

[12] Fudenberg, D., Levine, D. 1998. *The Theory of Learning in Games*. The MIT Press, Cambridge, MI.

[13] Harley, C. 1981. Learning the evolutionarily stable strategy. *Journal of Theoretical Biology*, 89, 611-633.

[14] Hsia, D. 1998. Learning in call markets. Department of Economics, University of Southern Califonria.

[15] McAllister, P. 1991. Adaptive approaches to stochastic programming. *Annals of Operations Research*, 30, 45-62.

[16] McKelvey, R., T. Palfrey. 1995. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10, 6-38.

[17] Milligan, G., M. Cooper. 1988. Study of variable standardization. *Journal of Classification*, 5, 181-204.

[18] Mookerjee, D., B. Sopher. 1994. Learning behavior in an experimental matching pennies game. *Games and Economic Behavior*, 7, 62-91.

[19] Mookerjee, D., B. Sopher. 1997. Learning and decision costs in experimental constant-sum games. *Games and Economic Behavior*, 19, 97-132.

[20] Morgan, J., M. Sefton. forthcoming. An experimental investigation of unprofitable games. *Games and Economic Behavior*.

[21] Nagel, R., F. Tang. 1998. Experimental results on the centipede game in normal form: An investigation on learning. *Journal of Mathematical Psychology*, 42, 356-384.

[22] Rapoport, A., W. Amaldoss. 2000. Mixed strategies and iterative eliminateion of strongly dominated strategies: An experimental investigation of states of knowledge. *Journal of Economic Behavior and Organization*, 42, 483-521.

[23] Robinson, J. 1951. An iterative method of solving a game, *Annals of Mathematics*, 54, 296-301.

[24] Roth, A. 1995. Introduction. In J. H. Kagel and A. E. Roth (eds.) *The Handbook of Experimental Economics*. Princeton: Princeton University Press.

[25] Roth, A., I. Erev. 1995. Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8, 164-212.

[26] Sarin, R., F. Vahid. 1997. Strategy similarity in learning. University of Texas, A&M Working Paper.

[27] Selten, R. forthcoming. Bounded rationality and learning. In E. Kalai (Eds.), Collected Volume of Nancy Schwartz Lectures. 1-13. Cambridge University Press.

[28] Selten, R., R. Stoecker. 1986. End behavior in sequences of finite prisoner's dilemma supergames: A learning theory approach. *Journal of Economic Behavior and Organization*, 7, 47-70.

[29] Stahl, D. 2000. Rule learning in symmetric normal-form games: Theory and evidence. *Games and Economic Behavior*. 32, 105-138.

[30] Van Huyck, J., J. Cook, R. Battalio. 1997. Adaptive behavior and coordination failure. *Journal of Economic Behavior and Organization*, 32, 483-503.

[31] Vriend, N. 1997. Will reasoning improve learning? *Economics Letters*, 55, 9-18.

Table 1: Payoffs in centipede games, Nagel-Tang (1998)

Even player number choices

		2	4	6	8	10	12	14
	1	4 1	4 1	4 1	4 1	4 1	4 1	4 1
Odd player number choices	3	2 5	8 2	8 2	8 2	8 2	8 2	8 2
	5	2 5	3 11	16 4	16 4	16 4	16 4	16 4
	7	2 5	3 11	6 22	32 8	32 8	32 8	32 8
	9	2 5	3 11	6 22	11 45	64 16	64 16	64 16
	11	2 5	3 11	6 22	11 45	22 90	128 32	128 32
	13	2 5	3 11	6 22	11 45	22 90	44 180	256 64

Table 2: Relative frequencies (%) choices in centipede games, Nagel-Tang (1998)

Odd numbers	%	Even numbers	%
1	.5	2	.9
3	1.6	4	1.7
5	5.4	6	11.3
7	26.1	8	33.1
9	33.1	10	31.1
11	22.5	12	14.3
13	10.8	14	7.7

Table 3: Transitions after lower-number (take) choices, Nagel-Tang (1998)

	Choice in period t+1 after “Take”							
choice in t	2	4	6	8	10	12	14	Total no.
2	.28	.08	0.32	0.08	0.12	0.04	0.08	25
4	.11	.11	0.40	0.15	0.15	0.06	0.02	47
6		.05	.32	0.41	0.14	0.06	0.01	296
8		.01	.05	.56	.36	.02	.01	594
10			.01	.12	.73	.14	.01	353
12			.03	.05	.07	.83	.02	59
	1	3	5	7	9	11	13	Total no.
1	.07	.29	.21	.07	.21	.07	.07	14
3	.04	.09	.44	.13	.18	.09	.02	45
5	.01	.06	.20	.47	.15	.08	.03	156
7		.01	.04	.60	.28	.07		617
9			.01	.08	.62	.26	.03	545
11					.17	.60	.23	173
13						.09	.91	46

Table 4: Transitions after higher-number (pass) choices, Nagel-Tang (1998)

	Choice in period t+1 after "Pass"							
choice in t	2	4	6	8	10	12	14	Total no.
2								0
4		.50			.50			2
6	.08	.23	.15	.33	.18	.03		39
8	.01	.04	.29	.49	.15	.04	.01	388
10	.01	.01	.08	.40	.40	.06	.03	572
12		.01	.03	.10	.21	.54	.11	364
14				.06	.10	.19	.65	231
	1	3	5	7	9	11	13	Total no.
3	1.00							1
5		.60	.20		.20			5
7	.01	.06	.25	.48	.10	.06	.04	156
9		.01	.04	.33	.48	.11	.02	446
11	.01		.02	.10	.31	.43	.12	490
13			.01	.05	.10	.34	.50	276

Table 5. Log Likelihoods and Parameter Estimates of the Various Adaptive Learning Models

Model		LL		Mean Parameter Estimates				
	Number of Parameters	In sample	Out of sample	ϕ	δ	κ	N(0)	λ
Odd Players								
Reinforcement	2	-2713.2	-1074.5	0.92	0.00	1.00	1.00	0.01
Belief	3	-3474.2	-1553.1	1.00	1.00	0.00	100.00	0.57
EWA, Recent Actual Payoff	5	-2667.6	-1069.8	0.91	0.14	1.00	1.00	0.01
EWA, Payoff Clairvoyance	5	-2596.6	-1016.8	0.91	0.32	1.00	1.00	0.01
Even Players								
Reinforcement	2	-2831.8	-991.7	0.92	0.00	1.00	1.00	0.01
Belief	3	-3668.9	-1556.0	0.87	1.00	0.00	0.16	0.04
EWA, Recent Actual Payoff	5	-2811.9	-983.0	0.91	0.15	1.00	1.00	0.01
EWA, Payoff Clairvoyance	5	-2791.4	-953.2	0.90	0.24	1.00	7.91	0.13

Table 6. A Comparison between Single Agent Model and Mean of Individual-Level Model: EWA-Clairvoyance

Model	LL		Mean Parameter Estimates (Std. Dev.)				
	In sample	Out of sample	ϕ	δ	κ	N0	λ
Odd Players							
Single Agent Model	-2596.8	-1016.8	0.91	0.32	1.00	1.00	0.01
Average of Individual Model	-76.5	-35.2	0.92	0.44	0.38	1.84	0.13
	(27.3)	(15.2)	(0.15)	(0.34)	(0.45)	(3.59)	(0.15)
Even Players							
Single Agent Model	-2791.4	-992.3	0.90	0.24	0.00	7.91	0.13
Average of Individual Model	-80.7	-31.2	0.90	0.47	0.34	3.94	0.17
	(25.2)	(14.5)	(0.15)	(0.33)	(0.46)	(5.66)	(0.18)

Table 7: A Cluster Analysis Using Individual Level Estimates

Mean Parameter Estimates (Std. Dev.)							
Odd Players				Even Players			
Number of subjects	ϕ	δ	κ	Number of subjects	ϕ	δ	κ
20	0.96	0.40	0.07	21	0.96	0.48	0.02
	(0.07)	(0.35)	(0.10)		(0.08)	(0.36)	(0.03)
10	0.82	0.51	0.99	9	0.76	0.44	0.98
	(0.20)	(0.33)	(0.01)		(0.17)	(0.27)	(0.02)

Figure 1: EWA's Model Parametric Space

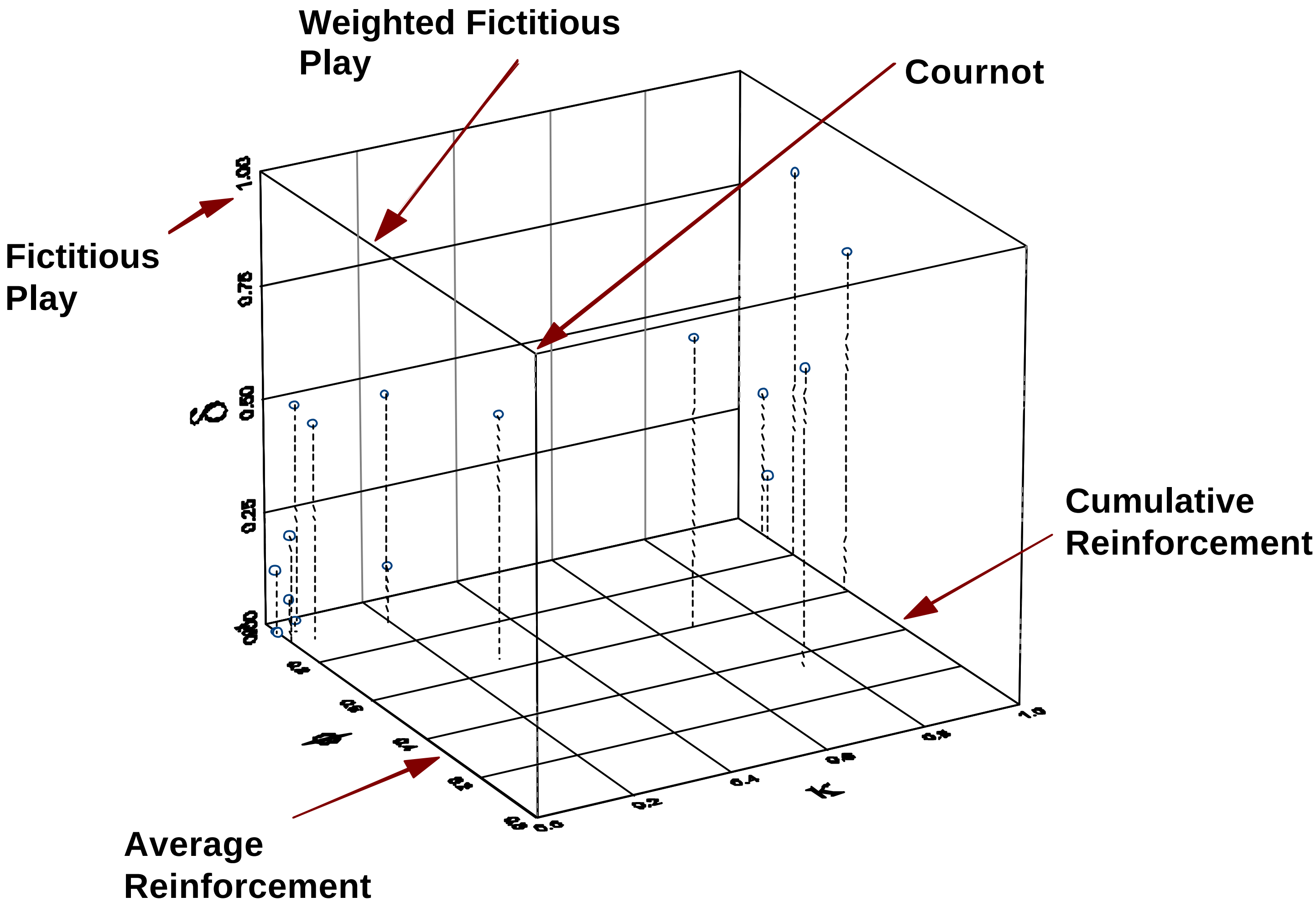

Figure 2a: Transition Behavior: Actual Data

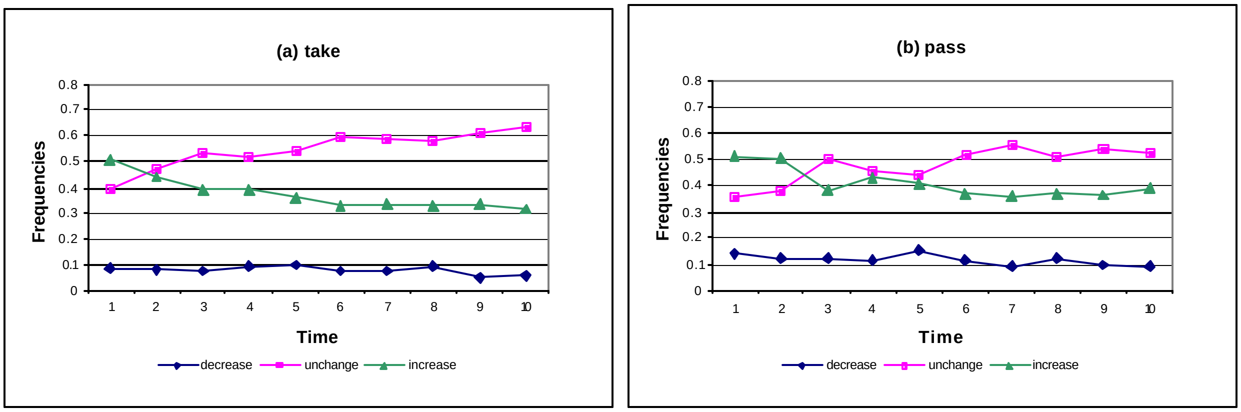

Figure 2b: Predicted Transition Behavior: EWA-Payoff Clairvoyance (Single Agent Model)

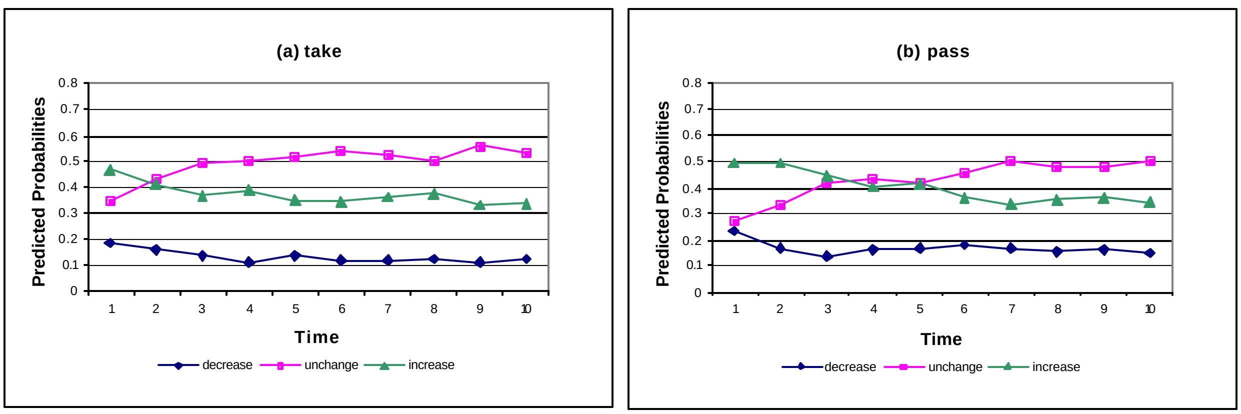

Figure 2c: Predicted Transition Behavior: EWA-Payoff Clairvoyance (Individual Model)

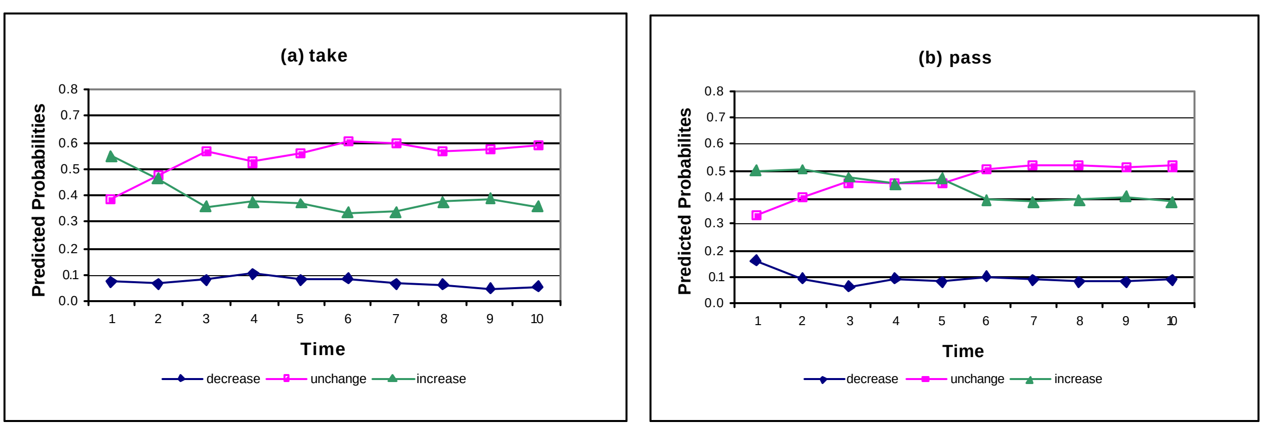

Figure 3a: EWA (Payoff Clairvoyance, Full Update) Model Parameter Patches: Odd Subjects

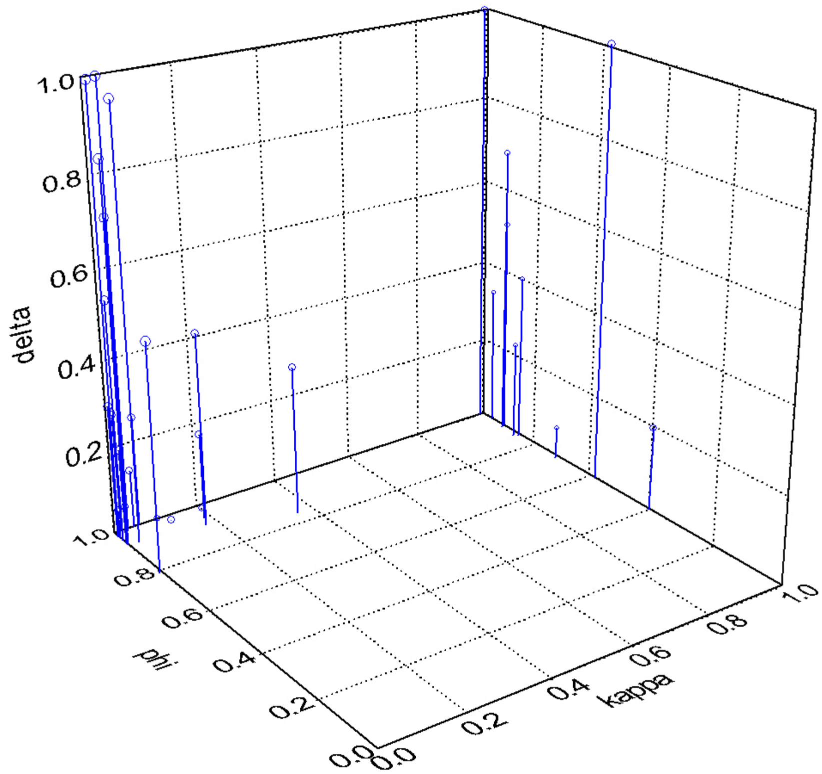

Figure 3b: EWA (Payoff Clairvoyance, Full Update) Model Parameter Patches: Even Subjects