Functional EWA:
A One-parameter Theory of Learning in Games

Teck H. Ho, Colin F. Camerer, and Juin-Kuan Chong[1]

November 7, 2002

Abstract

Functional experience weighted attraction (fEWA) is a one-parameter theory of learning in games. It approximates the free parameters in an earlier model (EWA) with functions of experience. The theory was originally tested on seven different games and compared to four other learning and equilibrium theories, then three more games were added. Generally fEWA or parameterized EWA predict best out-of-sample, but one kind of reinforcement learning predicts well in games with mixed-strategy equilibrium. Of the learning models, belief learning models fit worst but fit better than noisy (quantal response) equilibrium models. The economic value of a theory is measured by how much more subjects would have earned if they followed the theory's recommendations. Most learning theories add value (though equilibrium theories often subtract value) and fEWA and EWA usually add the most value.

[1]Ho: Haas School of Business, University of California, Berkeley, Berkeley, CA 94720-1900, hoteck@haas.berkeley.edu. Camerer: Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, camerer@hss.caltech.edu. Chong: National University of Singapore, 1 Business Link, Singapore 117951, bizcjk@nus.edu.sg. Direct correspondence to the first author. Thanks to participants in the 2000 Southern Economics Association meetings, the Wharton School Decision Processes Workshop, the University of Pittsburgh, the Berkeley Marketing Workshop, the Nobel Symposium on Behavioral and Experimental Economics (December 2001) and C. Mónica Capra, David Cooper, Vince Crawford, Ido Erev, Guillaume Frechette and two referees for comments.

"In nature hybrid species are usually sterile, but in science the reverse is often true"– Francis Crick (1988, p. 150)

The power of equilibrium models of behavior in strategic interactions comes from their ability to produce precise predictions using only the structure of a game and assumptions about rationality. Statistical models of learning should strive to be as parsimonious and precise, while also predicting the time path of actual observations more accurately than equilibrium theories can. Most learning models do this by specifying a formula for predicting future choices from past experiences (often at the population level), using one or more free parameters which are typically estimated from data. This paper describes a theory of learning in decisions and games called fEWA, with only one parameter. fEWA predicts the time path of individual behavior in any normal-form game as a function of previous history (given initial conditions). It is also easily extended to extensive-form games and games with incomplete information.

The key innovation in fEWA is the replacement of parameter values with functions of players' experience, which can vary across games, individuals, and time periods. Replacing parameters with functions kills two birds with one stone. The first bird is explaining why estimated model parameter values vary significantly across games (as earlier research showed). The functions in fEWA reproduce these cross-game differences endogenously, through the interaction between experience and game structure. The second bird is econometric parsimony. By replacing parameters with functions, only one free parameter needs to be estimated or fixed a priori. (The parameter captures sensitivity of players to differences in numerical ratings of strategies; it is essentially impossible to fit data well, and hence have a zero-parameter theory, without it.[2]) There are other one-parameter theories but they do not predict as well across games as fEWA does because best-fitting parameters tend to vary systematically across games.

fEWA was developed to fit and predict data from seven experimental data sets, and is compared to general versions of belief and reinforcement learning, and quantal response equilibrium. In out-of-sample forecasting, either fEWA or its parameterized precursor, EWA, tend to predict best, although a version of reinforcement predicts as well in some cross-game forecasting.

The paper makes three distinct contributions. First, the fEWA model is introduced to minimize the number of parameters which must be estimated (there is only one) and to allow parameters to flexibly self-adjust over time and across games.

[2]If the goal is to predict the most likely choice, fEWA can be reduced to a zero-parameter theory by setting the experience weight $N(0)$ to 0 (see below).

Second, the model is estimated on seven games and compared to four other models. fEWA is simpler than most other models and usually fits as well or better. In addition, because readers of the first version of this paper were concerned that the model was overfitted to the seven games we studied, we collected an (essentially) random sample of three new games and estimated all five models on those data too. This sort of statistical model competition is most useful when many different models are being tested on different sets of data, which is precisely the case currently in research on learning in game experiments.

Third, we introduce a new criterion for judging usefulness of theories– economic value. The economic value of a theory is measured by how well model forecasts of behavior of other players would improve a player's profitability if best responses to those forecasts were substituted for the player's actual choices. Most learning theories have positive economic value, and EWA or fEWA add the most economic value in most of the seven games we study.

Economists of all sorts should be interested in this paper because learning is important for economics. Laboratory control over information and incentives enables precise testing of which statistical models describe the path of learning best. Much as in physical sciences, in the lab we can see how theories perform in describing behavior (and giving valuable advice) cheaply before using them in more complex applications. While fEWA is crafted to explain learning in repeated games, sensible extensions of it could be applied to field settings such as evolution of economic institutions (e.g., internet auctions or pricing), investors and policymakers learning about equity market fluctuations or macroeconomic phenomena, and consumer choice. For example, a variant of the EWA theory is used by Ho and Chong (2002) to fit and predict 130,000 product choices by consumers. Their theory uses 80% *fewer* parameters than the leading theory used in marketing and predicts 20% better; it is also being used by supermarket chains to forecast sales. Readers who are interested in learning in field settings should be interested in how subjects learn in experimental games because understanding learning in the lab will surely help us understand learning in the field.

1 EWA learning and its limits

In earlier work, we proposed a model of learning called experience-weighted attraction (EWA) theory (Camerer and Ho 1998, 1999). Learning in EWA is characterized by changes in (unobserved) attractions based on experience. EWA was designed to be a gene-splice or hybrid of two models, reinforcement and belief learning, which have been used to study learning in games. The EWA model wraps a parametric skin around both of those theories, which are historically-interesting special cases on the boundary of the parameter space.

Attractions determine the probabilities of choosing different strategies through a logistic response function. For player i, there are m_i strategies (indexed by j) that have initial attractions denoted $A_i^j(0)$ (either estimated as free parameters from the data, specified by some theory of initial conditions, or "burned in" using the first period data). Denote i's j'th strategy by s_i^j, chosen strategies by i and other players (denoted $-i$) by $s_i(t)$ and $s_{-i}(t)$, and player i's payoffs by $\pi_i(s_i^j, s_{-i}(t))$. Define an indicator function $I(x, y)$ to be zero if $x \neq y$ and one if $x = y$. The EWA attraction updating equation is[3]

$$A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + [\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))}{N(t-1) \cdot \phi \cdot (1-\kappa) + 1} \tag{1.1}$$

and the experience weight is updated according to $N(t) = N(t-1) \cdot \phi(1-\kappa) + 1$.

The parameter δ is the weight placed on foregone payoffs. It presumably is affected by imagination (in psychological terms, the strength of counterfactual simulation) and reliability of information about foregone payoffs (Heller and Sarin, 2000). The parameter ϕ reflects decay of previous attractions due to forgetting or to deliberate ignorance of old experience when the learning environment is changing. The parameter κ controls the rate at which attractions grow. When $\kappa = 0$ attractions are weighted averages of reinforcements and decayed lagged attractions; when $\kappa = 1$ attractions cumulate. The growth rate of attractions is important because in the logit model the <u>difference</u> in attractions determines the spread of choice probabilities. The initial experience weight $N(0)$ is like a strength of prior beliefs and is estimated using data. Since it usually plays a minor role in predicting learning, we restrict $N(0) = 1$ in our specification of fEWA.[4]

A logit response function is used to map attractions into probabilities:

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}} \tag{1.2}$$

[3]This updating equation assumes that subjects know the payoffs of strategies that were not chosen. In Ho, Wang and Camerer (1999), we apply EWA model to games where such payoffs are not available by allowing subjects to learn about them through experience. See Chen and Khoroshilov (2000) for a similar extension.

[4]We switched notation in the denominator (previously we denoted the product $\phi \cdot (1-\kappa)$ by a single variable ρ), because using κ makes cumulation versus averaging more transparent.

where λ is the response sensitivity.

A key insight from our earlier work is that reinforcement and belief learning approaches are closely related in an interesting way.[5] When $\delta = 0$, the EWA rule is the same as models in which only chosen strategies are reinforced, originating in studies of animal learning. When $\kappa = 1$ the rule is a simpler form of cumulative reinforcement model studied by Harley (1981) and Roth and Erev (1995) (see also Bush and Mosteller, 1955; Cross, 1983; McAllister, 1991; Arthur, 1991). When $\kappa = 0$ the rule is like the averaging reinforcement model of Roth, Barron, Erev and Slonim (2002).

When $\delta = 1$ and $\kappa = 0$, the EWA rule is equivalent to belief learning using weighted fictitious play (Fudenberg and Levine, 1998). The EWA rule shows that belief learning is *not* fundamentally different than reinforcement. Instead, belief learning is a kind of reinforcement in which unchosen strategies are reinforced just s strongly as chosen ones, and attractions are averages of reinforcements rather than cumulations.

A graphical way to see the relation of different learning rules is a cube showing configurations of parameter values (see Figure 1). Each point in the cube is a triple of parameter values which specifies a precise updating equation (leaving aside λ and initial conditions). The cube shows the EWA family of learning rules. Corners and vertices of the EWA cube correspond to boundary special cases.

The corner of the cube with $\phi = \kappa = 0, \delta = 1$, is Cournot best-response dynamics. The corner $\kappa = 0, \phi = \delta = 1$, is standard fictitious play (Brown, 1951 and Robinson, 1951). The edge connecting these corners, $\delta = 1, \kappa = 0$, is the class of weighted fictitious play rules (e.g., Fudenberg and Levine, 1998). The edges with $\delta = 0$ and κ equal to zero or one are averaging and cumulative choice reinforcement rules.

The EWA cube is a visual aid to show the relations and differences among theories. But EWA is also a bet that the learning rules people actually use have parameter values which are in the interior of the cube rather than on vertices and corners. (That is, as Francis Crick suggested in the quote that opened this paper, a scientific hybrid may work better.) Reinforcement theories with $\delta = 0$ ignore foregone payoffs entirely.[6] Belief learning using weighted fictitious

[5] See also Cheung and Friedman, 1997, pp. 54-55; Fudenberg and Levine, 1998, pp. 1084-1085; Hopkins, in press.

[6] This assumption is implausible when foregone payoffs are known, and has been rejected by several studies comparing different information conditions, e.g., Mookerjhee and Sopher (1994), and Rapoport and Erev (1998);

play ($\delta = 1$) ignores the difference between received and foregone payoffs, which is also unlikely.[7] Put differently, reinforcement models assume that received payoffs matter more than foregone payoffs ($\delta < 1$) and belief learning says that foregone payoffs do matter ($\delta > 0$). Both intuitions are plausible. EWA allows them both if δ is between zero and one. Intermediate estimated values of δ could result if some subjects learn according to reinforcement and others according to weighted fictitious play, but direct tests allowing "latent class" heterogeneity show this is not so (Camerer and Ho, 1998).

Estimates by ourselves and others (see Camerer, Hsia, and Ho, in press, for a summary) have shown in 31 data sets that EWA generally fits (adjust for degrees of freedom) and predicts out-of-sample more accurately than the special cases of reinforcement and weighted fictitious play, except in games with mixed-strategy equilibrium (where all models only improve a little on Nash equilibrium, as we see below).

However, EWA is subject to two criticisms (which fEWA addresses). The first is that it has too many free parameters. Our empirical work anticipated this criticism by penalizing more complex theories for extra parameters, and measuring the ability of both simple and complex theories to forecast out-of-sample (where complex models have no automatic advantage). Nonetheless, having fewer parameters to estimate is often helpful. fEWA has only one, which is fewer than most other theories have.

The second criticism is that estimated parameter values vary across games.[8] Figure 1 also shows estimated EWA parameter triples from twenty games (see Camerer, Hsia, and Ho, in press). Each point corresponds to estimates from a different game. If one of the special case theories is a good approximation to how people generally behave across games, the parameters will cluster in the corner or vertex corresponding to that theory. In fact, parameters tend to be sprinkled around the cube. Estimates from coordination games usually have high values of δ and κ. Estimates from games with mixed equilibria tend to have low δ and κ (close to the averaging reinforcement corner with ϕ close to one). Roth, et al. (2002) note that our earlier work found "very different parameters in, apparently, very similar constant sum games. Their

and Van Huyck, Battalio and Rankin (1996).

[7]There is substantial evidence that people underweight opportunity costs compared to out-of-pocket costs (e.g., Kahneman, Knetsch, and Thaler, 1991). Since the difference between foregone and received payoffs is an opportunity cost (or gain), if it is underweighted then $\delta < 1$.

[8]Note that parametric variation across games is common when other learning models are estimated, e.g., Crawford (1995); Cheung and Friedman (1997), so it is not only a feature of the EWA estimates.

[Camerer and Ho's] research leads to the pessimistic conclusion that, at least currently, it is impossible to predict behavior in a new situation." fEWA is designed to meet this prediction challenge because it generates different parameter values from the interaction between specified functions and experience, and hence generates different parameters in different games.

2 fEWA

fEWA replaces the three central parameters of EWA, ϕ, δ, κ with deterministic functions $\phi_i(t)$, $\delta_i(t)$, $\kappa_i(t)$ of player i's experience up to period t. These functions determine parameter values for each player and period, which are then plugged into the EWA updating equation to determine attractions. Updated attractions determine choice probabilities according to the logit rule, given a value of λ. Standard methods for optimizing fit given λ can then be used to find which λ fits best.

2.1 The change-detector function $\phi_i(t)$

The decay rate ϕ is sometimes interpreted as forgetting, an interpretation carried over from reinforcement models of animal learning. Certainly forgetting does occur, but the more important variation in $\phi_i(t)$ across games is probably a player's perception of how quickly the learning environment is changing. The function $\phi_i(t)$ should therefore "detect change". As in physical change detectors (e.g., security systems or smoke alarms), the challenge is to detect change when it is really occurring, but not falsely mistake noise for change too often.

The core of the function is a "surprise index", the difference between the other players' strategies in the window of the last W periods and the average strategy of others in all previous periods (where W is the minimal support of Nash equilibria). We specify the function in terms of relative frequencies of strategies, without using information about how strategies are ordered, so it can be applied to non-ordered strategies (e.g., rows in a normal-form game). The change-detector function $\phi_i(t)$ is

$$\phi_i(t) = 1 - .5(\sum_{j=1}^{m_{-i}} [\frac{\sum_{\tau=t-W+1}^{t} I(s_{-i}^j, s_{-i}(\tau))}{W} - \frac{\sum_{\tau=1}^{t} I(s_{-i}^j, s_{-i}(\tau))}{t}]^2) \tag{2.1}$$

The term $\frac{\sum_{\tau=t-W+1}^{t} I(s_{-i}^j, s_{-i}(\tau))}{W}$ is the j-th element of a vector that simply counts how often strategy j was played by the others in periods $t-W+1$ to t, and divides by W. The term

$\frac{\sum_{\tau=1}^{t} I(s_{-i}^{j}, s_{-i}(\tau))}{t}$ is the relative frequency count of the j-th strategy over all t periods.[9] To measure change, we take the differences in corresponding elements of the two frequency vectors, square them, and sum over strategies. Since the maximum difference is two, the function is normalized by dividing the sum of squared differences by two, and subtracting the normalized figure from one. When recent observations of what others have done deviate a lot from all previous observations, the deviations in strategy frequencies will be high and ϕ will be low. When recent observations are like old observations, ϕ will be high.

While the change-detector was not derived explicitly from axioms, it was crafted to have some simple properties that are appealing. For example, the normalization simply ensures that ϕ is always (weakly) between zero and one. Another sensible property is that a long sample of previous experience should be used (i.e., ϕ close to one) when the environment is noisy. This requires ϕ to be larger when there is more dispersion in previous choices, which is guaranteed by squaring the deviations between current and previous history. Consider the limiting case where there is much experience. The period-t frequency vector is $(0, 0, \cdots, 1, \cdots, 0)$. The sum (across vector elements) of the squared deviations between this and any history vector $(1/f_1, 1/f_2, \cdots, 1/f_k)$ (where f_k represents previous frequency of choice k) is minimized when all frequencies are equal ($f_i = 1/k$), which gives the highest ϕ.[10]

Another property ϕ was designed to have is that it sticks close to one unless there is an unmistakably persistent change in what others are doing. It is dangerous to let ϕ become too low because doing so erases everything that has been learned, by giving a low weight to the previous attractions which summarize previous experience. The $\phi_i(t)$ function dips lowest in the extreme case in which one strategy is played until $t-1$, and then a surprise occurs and a new strategy is played. In that case, $\phi_i(t)$ is $\frac{2t-1}{t^2}$. This expression declines gracefully toward zero as the string of identical choices up to period t grows longer. (For t=2, 3, and 10 the $\phi_i(t)$ values are .75, .56, and .19.) This embodies the principle that a new choice is bigger surprise (and should have lower ϕ) if it follows more identical choices in a row. Another interesting special case is when different strategies have been played in every period up to $t-1$, and another different strategy is played in period t. (This is often true in games with large strategy spaces.) Then $\phi_i(t) = .5 + \frac{1}{2t}$, which starts at .75 and asymptotes at .5 as t increases.

[9] In the case of games with multiple players, frequency count of the relevant aggregate statistics is used. For example, in median action game, frequency count of the median strategy by all other players in each period is used.

[10] This example excludes the knife-edge case in which the previous frequency of the period-t choice is one, so the sum of squared deviations is zero and $\phi = 1$.

So far we have neglected an important detail: What's W? W is the smallest support of all the Nash equilibria (the number of strategies played with positive probability). In games with a pure strategy equilibrium, $W = 1$. In games with mixed equilibria W is larger than one. In these games, a certain amount of period to period change is expected. The number of strategies with positive probability, W, tells us roughly "how much" variation to expect, and hence, how many previous observations to average over to smooth perceived change.

2.2 Responsiveness $\delta_i(t)$

The parameter δ is the weight on foregone payoffs. We use $\delta_i(t) = \phi_i(t)/W$. Frankly, the specification $\delta = \frac{\phi}{W}$ is the feature of the fEWA model that is hardest to interpret or defend. It simply fits better than some other simple functions we tried, like setting $\delta = 1$ for better-responses or trying to track how responsive players were and adjusting their δ's accordingly. This specification embodies two separate properties– δ is tied to ϕ; and δ is decreasing in W.

Tying δ to the change measure ϕ recognizes the fact that best-responding to foregone payoffs is a good strategy when the environment is stable, so that δ should be near one when ϕ is near one. But when ϕ is low, the strategic environment is changing and information about past foregone payoffs is not likely to be a good guide to future choices. But then why should received payoffs be reinforced *relatively* more strongly than foregone payoffs (by a weight of one rather than low δ) when ϕ is low? There are two reasons. One is essentially econometric: When ϕ is low, then attractions from period $t-1$ are largely erased during the updating before period t. Reinforcing the chosen strategy from period t payoff with a weight of one partially "restores" information about what players are likely to do (since the erased lagged attractions and the previous choice which is strongly reinforced are likely to be correlated). The second reason is behavioral: Reinforcing chosen strategies more strongly than unchosen ones in low-ϕ environments models behavior of players who are especially likely to repeat what they did, like a "freezing" response to danger or "status quo bias", when the environment is changing.[11]

[11]Freezing is an immediate response to danger which is nearly universal across species (including humans, who instinctively stiffen when somebody sneaks up behind them and says "Boo!"). Freezing is presumably an adaptive response when predators are better at detecting movement than recognizing prey. Such a deep-seated response may lurk in the "old" or "animal" part of the human brain (the limbic system, which processes emotion and communicates with the prefrontal cortex that controls action; see LeDoux, 1996). Status quo bias refers to an exaggerated preference for the choice one has made in the past, even if the choice is assigned randomly (e.g., Samuelson and Zeckhauser, 1988; Kahneman, Knetsch, and Thaler, 1991). Experiments show that status quo

The reason why ϕ is divided by W to generate δ is that estimates of δ in games with mixed equilibria are generally lower than in games with pure equilibria. In fact, estimates of δ are often close to zero when the number of strategies in the mixed equilibrium is more than three or so. Dividing ϕ by W therefore creates values of δ which approximate those estimated in these games.

2.3 The exploitation parameter $\kappa_i(t)$

The parameter κ controls the growth rate of attractions. When $\kappa = 0$ attractions are weighted averages of lagged attractions and ($\delta-$weighted) payoffs, so (if initial attractions are scaled to payoffs) the attractions are bounded by payoffs. If $\kappa = 0$ attractions cannot grow too far apart. Fixing λ across periods, this means it is difficult to predict very sharp convergence in later periods (as we sometimes see in the lab). That's because using the logit probability function, the degree of sharpness or convergence in probability (i.e., the difference between the highest and lowest choice probabilities) depends only on the *difference* in attractions, which is multiplied by sensitivity parameter λ. When attractions are bounded by payoffs, attractions cannot grow too far apart so the $\lambda-$weighted differences cannot be too large. (This could be remedied by choosing a higher λ but that predicts behavior which is sharp early one, contrary to observations.) When $\kappa = 1$, however, attractions are (decayed) cumulations of previous (weighted) payoffs. Then attractions can grow larger and larger– they can be multiples of payoffs– and consequently, choice probabilities can grow further apart.

Psychologically, κ can be interpreted as the extent to which players "explore" by choosing different strategies, relative to how quickly they "exploit" what they have learned by switching to a constant choice of the strategy which has performed the best in the past.[12] Players with low κ are constantly exploring– they just keep track of average (δ-weighted) payoffs. When players "exploit" they commit to a strategy, even if its average previous payoff is not much larger than the average payoffs of other strategies. One way to model this is to let κ move toward one as players shift toward exploiting what they have learned. If payoffs are positive, a

bias is stronger when there are more sensible options available to switch to (loosely corresponding to W). This number-of-option effect can be captured in a model like ours by putting more reinforcement on the previous, status quo, choice, and putting reinforcement on alternative strategies which declines with W, precisely as in our model.

[12] The exploration- exploitation tradeoff is studied formally in the multi-armed bandit literature (Gittins, 1989), and is also of interest to computer scientists designing machines to learn, see Sutton and Barto (1998).

higher κ means players are basically rewarding a strategy they choose a lot, simply for being chosen (assuming $\delta < 1$). This is one way of characterizing lock-in empirically.

This line of argument suggests using variation in how frequently a player uses different strategies to track when they explore and when they exploit. We use the player's past behavior to tell us whether they explore or exploit and when they switch. The degree of exploration versus exploitation can be measured by the spread in probability of a player's observed choices. A standard measure of spread is the Gini coefficient, typically used to measure income inequality. We use the Gini coefficient too, where choice proportion is akin to income: When a player is exploring, the probabilistic 'income' will be spread to many strategies, and the Gini will be low. When a player has locked into one strategy, all her probability is allocated to that strategy and the Gini will be high (close to one).[13]

To calculate the Gini coefficient for subject i, first rank strategies from most-probable to least-probable (using observed choice frequencies). Denote the rank-ordered choice proportions of these strategies by $f_i^{(1)}(t)$ to $f_i^{(m_i)}(t)$. Then plot a cumulative probability distribution which measures the total probability of the strategies used as frequently as j or less frequently, $C_i(j,t) = \sum_{k=1}^{j} f_i^{(k)}(t)$. This calculation gives j points; use linear interpolation to create a piecewise-linear function connecting the points. The Gini coefficient is then the area between the identity line and the interpolated function passing through the $C_i(j,t)$ points, normalized so that Gini coefficients range from zero (when all strategies are played equally often) to 1 (when one strategy is played all the time).

The normalized Gini coefficient on strategy frequencies is then:

$$\kappa_i(t) = 1 - 2 \cdot \{\sum_{k=1}^{m_i} f_i^{(k)}(t) \cdot \frac{m_i - k}{m_i - 1}\} \tag{2.2}$$

where $f_i^k(t)$ are ranked from the lowest to the highest.[14]

This κ function reflects the following thought process: A player tracks her actual choice frequencies. When the spread is low, the player is still exploring and wants to keep attractions

[13] We also tried the sum of squared probabilities, a Herfindahl index often used to measure industrial concentration. This number is usually too low to fit well.

[14] For instance, in the median action game, suppose the relative choice frequencies for player i up to period t for actions 1-7 are 0, .0, .2, .4, .3, .0, and .1 respectively. Then we have $f_i^{(1)}(t) = f_i^{(2)}(t) = f_i^{(3)}(t) = 0.0$, $f_i^{(4)}(t) = 0.1$, $f_i^{(5)}(t) = 0.2$, $f_i^{(6)}(t) = 0.3$, and $f_i^{(7)}(t) = 0.4$.

Consequently, we have: $\kappa_i(t) = 1 - 2 \cdot \{0 + 0 + 0 + 0.1 \cdot \frac{7-4}{7-1} + 0.2 \cdot \frac{7-5}{7-1} + 0.3 \cdot \frac{7-6}{7-1} + 0\}$, which is $\frac{2}{3}$.

from cumulating, so she chooses a low κ so that attractions continue to be averages. However, as she learns and chooses one strategy more often, she begins to exploit what she has learned. Exploitation requires a way of guaranteeing that the most frequently-chosen strategies get chosen more and more often. One way to do this is to let attractions cumulate, so that frequently-chosen strategy attractions will grow larger and larger simply because they are chosen more often (assuming $\delta < 1$). Letting κ be a function of strategy "concentration" is one way to do this.

Using cumulation to capture exploitation of high-payoff strategies is related to other ideas. One may be familiar to economists– Polya urns, which have been used to explain economies from increasing returns (Arthur, 1989) A Polya urn starts with a distribution of balls (e.g., some red and some black). When a red ball is drawn, it is replaced, along with *another* red ball. Draws therefore generate a payoff *and* increase the chance that the same payoff will occur again. This is a simple model of increasing returns or learning-by- doing (drawing a red makes red more likely) with interesting mathematical properties.

The Gini coefficient captures a similar process. If one strategy is chosen often those choices lead to large κ, which means that strategy payoffs cumulate. When $\delta < 1$ (as is common), cumulation favors chosen strategies; so strategies which are chosen often get chosen more often in the future, as in the Polya urn.

2.4 Interpretation of the fEWA functions

The fEWA parameter functions are not grounded in familiar principles of rationality (like Bayesian updating). Since fEWA strives to outpredict models which *are* based on those principles, it is necessary to appeal to different principles. Furthermore, other specifications were tried (detailed in the previous draft of this paper, Ho, Camerer and Chong, 2001) and the ones described above fit substantially better than others. (Note that the sample of three new games we estimated after our earlier draft was written measures the extent to which the functions we chose might have overfit the seven games reported in the earlier draft.)

While the fEWA parameter functions are not derived from rationality principles-on purpose–, they can be thought of as procedurally rational (in Herbert Simon's language) because they are precise and are designed to accomplish a specific goal: Namely, to predict and perform well in a wide range of games. One can imagine a truly optimal learning rule which maximizes expected payoffs across a wide range of games.[15] However, we conjecture that such a rule would look more like fEWA than like other familiar rules. For example, fictitious play has good long-run properties in some environments but will not respond rapidly enough to changes in an environment. Cournot best-response changes too quickly in games with mixed equilibria. Weighted fictitious play is flexible enough to do well in both stationary environments and mixed games, but how does one pick the right weights, and let them self-adjust over time? fEWA does so automatically based on what is observed. Rather than derive a globally rational approach from axioms, our approach is like work in machine learning, which tries to develop robust heuristic algorithms which learn effectively in a wide variety of low-information environments (see Sutton and Barto 1998). Good machine learning rules are not provably optimal but perform well on tricky test cases and lifelike problems like those which good computerized robots could perform (navigating around obstacles, hill-climbing on rugged landscapes, difficult pattern recognition, and so forth).

The functions in fEWA have three advantages over many other theories.

First, fEWA is not simple in complexity (measured by lines of computer code to implement it, for example) but it *is* easy to estimate because it has only one free parameter, λ (which is hard to do without in empirical work[16]. The use of simple fictitious play and reinforcement theories in empirical analysis are often justified by the fact that they have few free parameters. By that criterion, fEWA should be used too.

A second advantage is that parameters in fEWA naturally vary across time, people, and games. In principle this variation might capture individual differences if they arise from experience. For example, some experiments on games with mixed-strategy equilibria allow subjects to

[15] See Josephson, 2001, who asks which parameter values emerge when players with different parameter values compete over time (a kind of "evolution of learning"). He generally finds that large δ values persist.

[16] It is conceivable that λ could also be specified ex ante but doing so will be difficult. The problem is that comparing values across games requires a standard unit of payoffs (we use dollars). However, changes in experimental currency which keep money earnings constant are likely to produce different behavior and require different values of λ (see McKelvey, Palfrey, and Weber, 2000). Furthermore, the model implicitly assumes that differences in strategy attraction calibrated in money terms drive differences in choice probability but other framing effects may matter. For example, if players are sensitive to percentage differences in payoffs rather than absolute differences, then using a fixed λ across games will not explain what they choose (they will act like they use a lower λ when a positive constant is added to payoffs). Pratt, Wise, and Zeckhauser (1979) show this effect using field data on price dispersion across product categories. If all these effects are eventually understood a theory of how λ varies across games could be developed but such a goal is ambitious and beyond the scope of this paper.

explicitly choose randomized strategies (see Camerer, 2002, chapter 2). In these experiments, some subjects play pure strategies and some play mixtures (e.g., Shachat, 2002). This difference in individual play is easily expressed by our $\kappa_i(t)$ function, which will tend toward one for purists and toward zero for mixers.

Because parameters can vary across time, a third advantage is that fEWA can mimic a very reduced form of "rule learning". Recall that different EWA parameter configurations correspond to specialized rules (such as cumulative choice reinforcement, fictitious play, or Cournot best-response dynamics). If parameters change throughout the game, those changes are like rule switching or rule learning, in which the rules players use change due to experience (as in Stahl, 1996, 1999, forthcoming, and Salmon, 1999). For example, if ϕ rises over time from 0 to 1, players are effectively switching from Cournot-type dynamics to fictitious play. If δ rises from 0 to 1, players are switching from reinforcement to belief learning. Rule learning is, of course, more general than the range of learning permissible in fEWA, and is an important competitor among those models which have many more parameters.

Robustness of physical materials and scientific models is usually defined as working well under a wide range of conditions (or, rarely "breaking"). One way to think about fEWA is that it achieves robustness by repair the main weaknesses in reinforcement and belief learning.

Reinforcement learning assumes players only use information about the payoffs they received, even when they know foregone payoffs. As a result, choice reinforcement sometimes underpredicts the rate of learning in games with large strategy spaces, in which players initially choose strategies in one part of the space then switch to strategies in an entirely different part of the space. The problem is that those strategies which are chosen late in the game were rarely picked early on so they were not reinforced (Camerer and Ho, 1998; and see the traveler's dilemma games below). Reinforcement also severely underpredicts learning in n-player games where most players earn no profits in a period and get no reinforcement (such as auctions or "winner-take-all" labor tournaments). This is evident below in beauty-contest games.[17]

Belief learning has a different weakness. Unlike reinforcement, belief models do learn quickly

[17]Of course, reinforcement learning can be speeded up if players reinforce payoffs relative to an aspiration level. But specifying aspiration levels requires two parameters– an initial aspiration level and an adjustment rate. EWA generates aspiration-based reinforcement with no extra parameters: Strategies only increase in probability (holding their lagged attractions constant) if their $\delta-$weighted payoffs are above the average $\delta-$weighted payoff. Thus, the δ-weighted payoff is *is* an aspiration level, which evolves endogenously over time without requiring any free parameters.

in games with shifting support and with many zero-payoff players, but belief learning gives no ready explanation for why the decay rates on previous observations differ across games. For example, old observations are decayed more rapidly in the continental divide and beauty contest games than in games with mixed equilibria and patent race games. Confronted with a brand new game, belief learning theories have no built-in way of guessing whether Cournot-like responsive dynamics ($\phi = 0$) or standard fictitious play ($\phi = 1$) will predict best. fEWA can do better by automatically adjusting ϕ as a function of initial conditions and game structure.

In the form we use, the fEWA model does require information about initial conditions (i.e, relative frequencies of first-period play) and information about the structure of the game– namely, the minimal support of Nash equilibria (W). As a practical matter, pinning down W boils down to guessing whether a game has a pure-strategy equilibrium, and whether it is symmetric or not, or has only mixed equilibria (and if so, how many strategies are used in the mixture). Even in field applications where the game is not controlled as in the lab, guessing whether W is one or is much larger is not hard to do.

3 fEWA predictions within and across games

In this section we compare in-sample fit and out-of-sample predictive accuracy of different learning models when parameters are freely estimated, and check whether fEWA functions can produce game-specific parameters similar to estimated values.

We use seven games: Games with unique mixed strategy equilibrium (Mookerjhee and Sopher, 1997); R&D patent race games (Rapoport and Amaldoss, 2000); a median-action order statistic coordination game with several players (Van Huyck, Battalio, and Beil, 1990); a continental-divide coordination game, in which convergence behavior is extremely sensitive to initial conditions (Van Huyck, Cook, and Battalio, 1997); a coordination game about entry to two markets of different sizes (Amaldoss and Ho, 2001); dominance-solvable p-beauty contests (Ho, Camerer, and Weigelt, 1998); and a traveler's dilemma game (Capra, Goeree, Gomez and Holt, 1999). Table 1 summarizes features of these games and the data. Three of the games are described in detail below.[18] Since one of our goals is to see whether fEWA can explain cross-

[18]The other four games are: Mixed-equilibrium games studied by Mookerjhee and Sopher (1997) which have four or six strategies, one of which is weakly-dominated; the nine-player median-action game studied by Van Huyck et al. (1990), in which players choose integer strategies 1-7 and earn payoffs increasing linearly in the

game variation in model parameters, we sample different classes of games. Sampling widely is also a good way to test robustness of any model of learning or equilibrium.[19]

3.1 Estimation method

The estimation procedure for fEWA is sketched briefly here and detailed in Appendix 7. Consider a game where N subjects play T rounds. For a given player i, the likelihood function of observing a choice history of $\{s_i(1), s_i(2), \ldots, s_i(T-1), s_i(T)\}$ is given by:

$$\Pi_{t=1}^{T} P_i^{s_i(t)}(t). \tag{3.1}$$

The joint likelihood function L of observing all players' choice is given by

$$L(\lambda) = \Pi_i^N \{\Pi_{t=1}^{T} P_i^{s_i(t)}(t)\} \tag{3.2}$$

We "burn in" the model by choosing the initial attractions $A_i^j(0)$ (the same for all i) that correspond to choice probabilities that match the actual population frequency of choices in the first period (given the estimate of λ).[20] (When data on initial choices are unavailable some theory of initial play could be used instead.[21]) Details of the "burn-in" are given in Appendix 7.1. The initial parameter values are $\phi_i(0) = \kappa_i(0) = .5$ and $\delta_i(0) = \phi_i(0)/W$. These initial values are averaged with period-specific values determined by the functions, weighting the initial value by $\frac{1}{t}$ and the functional value by $\frac{t-1}{t}$.

group median and decreasing linearly in the squared deviation from the median; dominance-solvable p-beauty contest games in which players choose numbers from 0 to 100 and the player closest to p times the average earns a fixed prize (for p equal to .7 or .9); and a coordination game in which n players simultaneously enter a large or small market and earn $2n$ (n) divided by the number of entrants if they enter the large (small) market.

[19]Another approach is to sample randomly within a class of games, although results are likely to be sensitive to which class of games is chosen.

[20]This is a small departure from some of our earlier work in which the $A_i^j(0)$ are estimated as free parameters. Estimation is infeasible in some of the games we study because there are many strategies (e.g., integer prices from 80 to 200) and we were reluctant to impose ad hoc functional forms to generate a parsimonious $A_i^j(0)$ distribution.

[21]A mixture of random behavior and "level-1" reasoning– best-responses to the belief that others will behave randomly– will generally be a good guess about what players would do in the first period (see Haruvy and Stahl, 1998). In the games we study, for example, level-1 behavior predicts choices of 35 in beauty contests, 7 in continental- divide games, 4 in median-action games, the large pot in entry-choice games, 5 and 4 in patent-race games for strong and weak players, and $200 - 2R$ in traveler's dilemma games. Combining these guesses with random initial behavior gives a good approximation to what players actually do in the first period.

Given the initial attractions and initial parameter values, attractions are updated using the EWA formula. fEWA parameters are then updated according to the functions above. Maximum likelihood estimation is used to find the best-fitting value of λ (and other parameters, for the other models) using data from the first 70% of the subjects. Then the value of λ is frozen and used to forecast behavior of the entire path of the remaining 30% of the subjects.[22] Payoffs were all converted to dollars (which is important for cross-game forecasting).

In addition to fEWA, we estimated the EWA model in Camerer and Ho (1999) and versions of belief-based (weighted fictitious play) and reinforcement models.[23] To put the models on a more even footing, we did *not* force the belief model to have initial attractions which are consistent with a common initial belief (as in our earlier work); we simply burned in the first-period data as for the other models.[24] We fit three versions of reinforcement ($\delta = 0$). Two versions included an experience weight $N(0)$ (which our 1999 paper did not) and fixed κ to be either zero or one (the latter is a simplified form of the model in Erev and Roth (1998)). A third version, which is quite different, is the two-parameter model used by Erev, Bereby-Meyer, and Roth (1999) and Roth et al. (2002). Their new approach sets $\phi = 1$ and $\kappa = 0$, updates only chosen strategies, uses logit probability instead of power, and divides attractions by a measure of payoff variability (see Appendix 7.3 for details). We report only results from this latest payoff-variability (PV) reinforcement model but performance of the earlier reinforcement models is similar. We also fit the one-parameter quantal response equilibrium (QRE) model (McKelvey and Palfrey, 1995, see Appendix 7.2 for details) as a static benchmark, which is tougher competition than Nash equilibrium.

[22]This is another departure from our earlier work, in which we used the first 70% of the observations from each subject, then forecasted the last 30%. We also tried our earlier method, and a hybrid in which the holdout sample consisted of both later periods for some subjects, and the entire path for new subjects. The results from the two different methods are not interestingly different.

[23]For simplicity, we ignore two other interesting approaches to individual learning– rule learning or "learning to learn" (e.g., Stahl 1999; Salmon, 1999); and "direction learning" (Selten and Stoecker, 1986). See Camerer (2002, chapter 6) for more details.

[24]This switch helps belief models a lot in some games. For example, in the Mookerjhee-Sopher games with mixed equilibrium one strategy is only weakly dominated, and is rarely chosen. Most prior belief specifications will assign an expected payoff to that strategy which is only a little less than the expected payoffs of undominated strategies and given stochastic response, will overpredict how often the dominated strategy is chosen.

3.2 Model fit and predictive accuracy

The first question we answer is how well models fit and predict on a game-by-game basis (i.e., parameters are estimated separately for each game). To guard against overfitting we estimate parameters using 70% of the subjects (in-sample calibration) and use those estimates to predict choices by the remaining 30% (out-of-sample validation). For in-sample estimation we report both hit rates (the fraction of choices predicted to be most likely which are actually picked) and a Bayesian information criterion (BIC) which subtracts a penalty $\frac{k \cdot ln(NT)}{2}$ from the LL. (Note that the BIC imposes a stiffer penalty than other information criteria like Akaike.[25]) For out-of-sample validation we report hit rates and LL.

Table 2 shows the results. The best fits for each game and criterion are printed in bold; hit rates which are less than the best but are statistically indistinguishable (by the McNemar test) are underlined. Across games, EWA is better or as good as all other theories judging by hit rate, and fits better according to BIC or LL in four of seven games. fEWA also has higher or equal hit rates than other models in most games. Reinforcement with PV has the best BIC and LL in two games. Of the learning models, belief learning fits worst (it never has the best BIC or LL and is only best on hit rate in forecasting the median-action game). QRE fits worst of all, except in games with mixed equilibria where most models are about equally good.

The bottom line of each panel in Table 2, labeled "pooled", shows results when a single set of common parameters is estimated for all games (except for game-specific λ). If fEWA is capturing parameter differences across games effectively, it should predict especially accurately, compared to other models, when games are pooled. Indeed, fEWA fits and predicts best by both criteria when data are pooled.

A tough test of robustness is to estimate all parameters on six of the seven games and use those parameters to predict choices in the remaining game for all subjects, for each of the seven games. Cross-game prediction has been used by others but only within a narrow class of games (2x2 games with mixed equilibria, Erev and Roth, 1998; and 5x5 symmetric games, Stahl, forthcoming). Our results test whether fitting a model on a coordination game, say, can

[25] Myung (2000) discusses model selection, including some recent measures which penalize theories for the flexibility of their functional form as well as for number of free parameters. He notes in an example that the squared deviation or squared error criterion (MSD, or MSE) is the measure which penalizes complex theories the *least* effectively. A very sensible measure, Bayesian model selection (BMS) reduces to the BIC when the modeler has a diffuse prior over parameter values.

predict behavior in a game with mixed equilibrium. Table 3 reports results from this kind of cross-game prediction. fEWA has the highest hit rate in four games; EWA is highest in two other games. Reinforcement with PV also predicts across games reasonably well; it has the best *LL* in three games. The biggest losers are belief models and QRE, which are usually much lower than the other models by any criterion.

The point of fEWA is to use only structural features of games and players' experience to create parameter values which are close to the EWA estimates across games. Figure 2 shows how well fEWA functional values of δ and ϕ match the estimates from EWA. Each pair of connected points represents one of the seven games and the pooled estimates. Open (closed) circles are EWA estimates (fEWA functional values). If the two are close together within each game the chords connecting points should be short, and if they are different across games they should be sprinkled around the square. The chords are short in about half the games (most of the long-chord deviation between the function averages and estimate are on the ϕ dimension rather than δ). The correlation of the parameter estimates and functional averages across the seven games is .92 for δ and .78 for ϕ.[26] The pooled estimates only differ by .01 and .02, respectively. Details are reported in Table A.1 in Appendix 7.4 (along with estimates for other models and standard errors).

In addition, Table A.2 in Appendix 7.4 shows how much fEWA functional values vary across time periods and across people. The variation is usually not very large; the interquartile range is typically from zero to .10. An interesting exception is κ in the three games with mixed equilibria. The interquartile range for average κ within subject (i.e., averaged across periods for each subject) is .20 or more in these games. That means some subjects are roughly choosing pure strategies while others are mixing across all strategies which shows the potential for the fEWA approach to detect individual differences.

Next we will show predicted and relative frequencies for three games. Corresponding graphs for *all* games can be seen at http://www.bschool.nus.edu.sg /depart/mk/bizcjk/fewa.htm. We chose these three games because each has interesting differences visible to the naked eye and each is representative of a different class– one has a unique mixed-strategy equilibrium, one has multiple Pareto-ranked pure equilibria, and one is dominance-solvable.

[26]The correspondence is much worse for κ, which is basically estimated to be either zero or one in EWA but only varies from around .4 to .8 in fEWA.

3.3 Games with unique mixed strategy equilibrium: Patent race

In the patent race game (Rapoport and Amaldoss, 2000), two players, one strong and one weak, are endowed with resources and compete in a patent race. The strong (weak) player has an endowment of 5 (4) and can invest an integer amount from zero to their endowments. Players invest simultaneously. They earn 10 minus their investment if their investment is strictly largest, and lose their investment if it is less than or equal to the other player's.

The game has an interesting strategic structure. The strong player can guarantee a payoff of five by investing the entire endowment (outspending the weak player), which strictly dominates investing zero. Eliminating the strong player's dominated (zero) strategy then makes investing one dominated for the weak player. Iterating in this way, both players can delete three strategies by iterated application of strict dominance. There is a unique mixed equilibrium in which strong (weak) players invest 5 (0) 60% of the time and play their other two (serially) undominated strategies 20% of the time.

Thirty six pairs of subjects played the game in a random matching protocol 160 times (with the role switched after 80 rounds); the 36 pairs are divided into 2 groups where random matching occurs within group. Choice frequencies do not change visibly across time so we plot frequencies of transitions between period $t-1$ and period t strategies instead, for strong players, using the within-game estimation and pooling across all subjects. (Weak player results are similar.) Figures 3a-f show the empirical transition matrix and predicted transition frequencies for five models on strong players. The key features of the data are a lot of transitions from 5 to 5, almost 40%, and roughly equal numbers of transitions (about 5%) from 1 to 1, and from 1 to 5 or vice versa.

Two models are clearly inferior: QRE does not predict differences in transitions at all (it is a benchmark, not a learning theory); and the belief-based model predicts too few 5-to-5 and 1-to-1 transitions. (Table 2 confirms that belief learning fits relatively poorly here.) Where does belief learning go wrong? Since belief learning assumes full responsiveness to foregone payoffs, it will often predict that players should move away from chosen strategies which were winners, if other strategies would have been even better. Note that (as the equilibrium predicts) weak players abandon hope and invest zero about half the time. As a result, when strong players invest 5, half the time they earn a payoff of 5 but they could have earned more by investing less (because the weak player invested nothing). Belief models therefore predict more switching away from investing 5 than is evident in the data. Both EWA and reinforcement approaches can explain the infrequency of transitions by multiplying the higher foregone payoffs, in the case where the strong player invests 5 and the weak player invests nothing, by δ. A low value of δ (estimated to be .36 in EWA and .31 in fEWA) therefore characterizes the sluggishness in switching and avoids the predictive mistake inherent in belief learning.

3.4 Games with multiple pure strategy equilibria: Continental divide game

Van Huyck et al. (1997) studied a coordination game with multiple equilibria and extreme sensitivity to initial conditions, which we call the continental divide game (CDG). The payoffs in the game are shown in Table 4. Subjects play in cohorts of seven people. Subjects choose an integer from 1 to 14, and their payoff depends on their own choice and on the median choice of all seven players.

The payoff matrix is constructed so that there are two pure equilibria (at 3 and 12) which are Pareto-ranked (12 is the better one). Best responses to different medians are in bold. The best-response correspondence bifurcates in the middle: If the median starts at 7 virtually any sort of learning dynamics will lead players toward the equilibrium at 3. If the median starts at 8 or above, however, learning will eventually converge to an equilibrium of 12. Both equilibrium payoffs are shown in bold italics. The payoff at 3 is about half as much as at 12. This game captures the possibility of extreme sensitivity to initial conditions (or path-dependence).

Their experiment used 10 cohorts of seven subjects each, playing for 15 periods. At the end of each period subjects learned the median, and played again with the same group in a partner protocol. Payoffs were the amounts in the table, in pennies.

Figures 4a-f show empirical frequencies (pooling all subjects) and model predictions. The key features of the data are: Bifurcation over time from choices in the middle of the range (5-10) to the extremes, near the equilibria at 3 and 12; and late-period choices are more clustered around 12 than around 3. (Figure 4a disguises the extreme path-dependence: Groups which had first-period medians below (above) 7 *always* converged toward the low (high) equilibrium.) Notice also that strategies 1-4 are never chosen in early periods, but are frequently chosen in later periods; and notice that strategies 7-9 are frequently chosen in early periods but never chosen in later periods. A good model should be able to capture these subtle effects by "accelerating" low choices quickly (going from zero to frequent choices in a few periods) and "braking" midrange choices quickly (going from frequent choices to zero).

QRE fits poorly because it predicts no movement.

Belief learning does not reproduce the asymmetry between sharp convergence to the high equilibrium and flatter convergence around the low equilibrium, because of a subtle weakness in belief learning. Note from Table 4 that the payoff gradients around the equilibria at 3 and 12 are exactly the same– choosing one number too high or low "costs" \$.02; choosing two numbers too high or low costs \$.08, and so forth. Belief learning computes expected payoffs, and the logit rule means only differences in expected payoffs influence choice probability. The fact that the payoff gradients are the same therefore means the spread of probability around the two equilibria must be the same; so belief learning predicts similar probability distributions around the high and low equilibria.

Then how do the EWA and reinforcement models generate the asymmetry? The trick is $\delta < 1$. At the high equilibrium, the payoffs are larger and so the difference between the received payoff and δ times the foregone payoff will be larger than at the low equilibrium if $\delta < 1$. This explains the sharper convergence around 12.[27]

Reinforcement with PV fits reasonably well, except it predicts frequencies for choices 3-5 which are too low (10% instead of 15%) and it predicts substantial early choice of strategies 1-2 which declines over time. (The EWA models smear a little probability at 1-2, and grow over time, to avoid a large likelihood penalty from missing the rare choices of 1 and 2 which come in later periods.)

3.5 Games with dominance-solvable pure strategy equilibrium: Traveler's dilemma

Capra et al. (1999) studied a dominance-solvable "traveler's dilemma" (introduced by Basu, 1984) in which two players must choose a number or 'claim' between 80 and 200. If the claims are equal, each player receives the amount claimed. If the claims are unequal, each of them receives the lower of the two claims. In addition, the person who makes the lower claim receives

[27]Numerically, a player who chooses 3 when the median is 3 earns \$.60 and has a foregone payoff from 2 or 4 of $\$.58 \cdot \delta$. The corresponding figures for a player choosing 12 are \$1.12 and $\$1.10 \cdot \delta$. The differences in received and foregone payoffs around 12 and around 3 are the same when $\delta = 1$ but the difference around 12 grows larger as δ falls (for example, for the fEWA estimate $\hat{\delta} = .69$, the differences are \$.20 and \$.36 for 3 and 12.) Cumulating payoffs rather than averaging them contributes to explaining the difference by "blowing up" the expected payoff differences over time.

a bonus R and the person who makes the higher claim pays a penalty of R. Let these claims be x_1 and x_2 respectively. Formally, the payoff to each player i is defined as follows:

$$\pi_i(x_i, x_{-i}) = \begin{cases} x_i & \text{if } x_i = x_{-i} \\ x_i + R & \text{if } x_i < x_{-i} \\ x_{-i} - R & \text{if } x_i > x_{-i} \end{cases} \tag{3.3}$$

The game is like one of imperfect competition in which two sellers both sell products at the lowest price (due to consumer shopping or "meet-or-release" clauses) and the seller who names the lowest price earns a goodwill reward while the high-price sellers suffers a reputational loss. The Nash equilibrium predicts that the lowest possible claim of 80 will be chosen by both players. This prediction is also insensitive to R. Their experiment used six groups of 9-12 subjects. The reward/penalty R was varied at 6 levels (5, 10, 20, 25, 50, 80). Each subject played 10 times (and played with a different R for five more rounds; we use only the first 10 rounds).[28] Figures 5a-f show empirical frequencies and model fits for R=50. (Other values of R give roughly similar results although R=50 illustrates differences across models best.) A wide range of prices are named in the first round. Prices gradually fall, between 91-100 in rounds 3-5, 81-90 in rounds 5-6, and toward the equilibrium of 80 in later rounds.

QRE predicts a spike at the equilibrium of 80. As λ rises, the QRE equilibria move sharply from smearing probability throughout the price range to a sharp spike at the equilibrium; there is no intermediate value of λ which can explain the combination of initial dispersion and sharp convergence in later periods.

The belief-based model predicts the correct direction of convergence, but overpredicts numbers in the interval 81-90 and underpredicts choices of precisely 80. The problem is that the incentive in the traveler's dilemma is to undercut the other player's price by as little as possible. Players only choose 80 frequently in the last couple of periods; before those periods it pays to choose higher numbers. The belief model also estimates $\hat{\phi} = .85$ and does not allow payoffs to cumulate, so there is a large burden of historical information which keeps the model from reacting quickly to frequent choices of 80 which come late in the game. EWA models explain the sharp convergence in late periods by cumulating payoffs and estimating $\delta = .63$ (for fEWA). Consider two players who choose 80 and 90 when R=50. The player who chose 80 is reinforced by 130, and choosing the best response 89 is reinforced by $139 \cdot .63$, or 88.1, so choosing 80

[28] We did not use $R = 10$ with 9 subjects, where there is always one subject gone unmatched in each round, to avoid making ad hoc assumptions on the learning behavior of the unmatched subjects.

is more strongly reinforced and she is likely to repeat that choice. In belief learning, the best response of 89 is reinforced by 139 so she is predicted to move *away* from 80, contrary to what happens.

The reinforcement model has a reasonable hit rate because the highest spikes in the graph often correspond with spikes in the data, but predicted learning is clearly more sluggish than in the data (i.e., the spikes are not high enough). Because effectively $\phi = 1$ and players are not predicted to move toward ex-post best responses, the model cannot explain why players learn to choose 80 so rapidly. (Reinforcement models with variable ϕ and no payoff variability term predict much better when R=50.)

4 The economic value of learning models

Most criteria used to judge fit and predictive accuracy of models are purely statistical, or are roughly equivalent to familiar statistics. But economic applications of theories demand a financial measure of what good theories are worth.

In this section we measure the economic value of a learning theory. Camerer and Ho (2001) defined a theory's economic value as the increase in a subject's profit from substituting learning theory recommendations– i.e., best responses based on the theory's prediction about what others will do– for their actual choices. This definition treats a theory as similar to the advice service professionals like consultants sell. The value of the advice is the difference in economic quality of the client's decisions with and without the advice.

To measure economic value, we use model parameters and a player's observed experience through period t to generate model predictions about what others will do in $t + 1$. That prediction is a probability distribution over choices by others; optimizing using this predicted distribution produces a choice with the highest expected value. We then compare the profit from making that choice in $t + 1$ (given what other players actually did in $t + 1$) with profit from the target player's actual choice.

Economic value is a good measure because it uses the distribution of predictions about what other players are likely to do, *and* the economic impact of those possible choices. A model could be statistically accurate, but not improve expected profits much, and will therefore have a low economic value. Oppositely, a model with a small improvement in statistical accuracy could have substantial economic value if the payoff landscape is "rugged" (sharply-peaked). An equilibrium theory that is correct has little economic value, by definition, since its recommendations will match the choices subjects make.

We use two methods to estimate model parameters and compute model recommendations: Using in-game estimated parameters (see Appendix Table A.1); and using estimates from the six other games to compute economic value in the seventh game, for each of the seven games (as in Table 3). Using all the data from a game to give advice to subjects in that game is like advising a client who has a large sample of experience in a particular situation for the analyst to estimate parameters with. Using only data from other games is like advising a client in a new situation who has no direct experience for the analyst to use to estimate parameters. In practice, economic value would generally fall between the bounds of economic value computed these two ways.

Table 5 shows the overall economic value– the percentage improvement (or decline) in payoffs of subjects from following a model recommendation rather than their actual choices. Most models have positive economic value in most games. Value generated from in-game estimates is usually higher than when parameters are estimated from different games.

Using either method of parameter estimation, either fEWA or EWA have the most economic value in five to seven of the seven games. Subject-by-subject analysis further show that EWA and fEWA (weakly) add value for a large majority of subjects[29], from 48% to 97% for out-of-game estimates.

In some games, the percentage improvement from economic value is small because the difference between observed payoffs and ex-post optimal payoffs– the payoffs that a perfectly clairvoyant player would have earned– is low, which puts an upper bound on economic value. (Any theory's economic value is bounded by the value of a clairvoyant theory.) In the continental divide game, for example, perfect forecasts would have improved profits by only 6.58% and fEWA improved profits (out-of-game) by 4.98%. So the fEWA forecast improves profits by only a small percentage, but achieves about 80% of the largest possible improvement.

Belief learning has positive economic value in most games. Reinforcement actually has negative economic value in one to three games. In continental divide games, reinforcement

[29]See our working paper at www.bschool.nus.edu.sg/depart/mk/bizcjk/fewa.htm for details. In some games, like coordination games which equilibrate quickly, estimates of economic value are often zero because subjects make precisely the same choices the learning theories recommend.

tends to underpredict how rapidly players move away from middle strategies. It therefore advises players to switch to low or high strategies less quickly than they actually do, which turns out to be bad advice. The beauty contest game is similar. Reinforcement predicts hardly any learning by other players, so it never advises players to chose numbers which are low enough.

The way economic value is computed in the analyses just described does not control for the boomerang effect of a recommended choice in period t on future behavior by others in periods $t+1$ and beyond. As a result, it is possible that a subject who adopted a theory's recommendation would make choices that would reduce long-run payoffs. (In coordination games of common interest, the boomerang effect could also be positive by teaching others to expect behavior which is mutually-beneficial.) Choices which are economically-valuable in the short run will also distort what others do in the future which may alter the statistically accuracy of estimated parameters, as in the Lucas critique of macroeconomic forecasting. The boomerang and Lucas-critique effects will be small in some of the games we study. In the coordination and p-beauty contest games the medians and averages are computed for groups of 7-9 players, so substituting one player's theoretically-optimal choice for the actual choice will not affect the median or average much, and hence does not affect future behavior much. In the traveler's dilemma players are randomly rematched so changing one player's current choice may affect her current partner's future behavior, but she is unlikely to be rematched with same partner again.

The ideal way to adjust for both the boomerang and Lucas-critique effects is to run experiments in which subjects who are programmed with updated learning rules, and choose best responses, compete with actual human subjects. Such a study is technically difficult[30] and beyond the scope of this paper.

A proxy is to divide actual subject choices into those that are consistent with the model advice and those that are not, and see which choices earn more. If the advice is good, the subjects who "coincidentally" followed the advice should earn more.

The comparison we make is analogous to a common consulting practice that divides competing firms in an industry into groups based on their organizational practices and then determines whether one group is superior to the rest, and hence which organizational practice is associated

[30]The technical difficulty is that best responses must be computed by computerized subjects as rapidly as human responses, which is challenging even with modern computing. Any integrated man-machine design also raises a question of what human subjects should be told about who they are playing with.

with high performance. Table 6 shows the results. The results are divided into five triples of columns, one triple for each learning model. The first column in each triple shows the difference between average payoffs of choices that are consistent with the model advice and of those that are not. (If the model advice is good these differences should be positive.) The second column in each panel is the p-value of the t-statistic testing whether the average payoffs from advice-consistent and advice-inconsistent choices are the same. The third column is the percentage of actual choices that are consistent with model advice.[31] Except for beauty contests (where advice-consistent choices are rare), there are large enough samples of advice-consistent and -inconsistent choices to permit powerful tests.

The crucial question is whether conclusions about economic value from the Table 5 measure– is economic value positive?, and how does it vary across theories?– are replicated when this coincidence measure is used. Most conclusions are the same using either method.

The learning theories add comparable amounts of positive economic value using the coincidence measure. Of the learning theories, reinforcement adds less value in three games (and subtracts value, insignificantly, in beauty contests) but adds the most value in pot games. QRE adds the least value in every game but one, and subtracts value in three of seven games.

5 Robustness of fEWA

An obvious concern about fEWA is that the specification search for the three parametric functions $(\phi_i(t), \delta_i(t), \kappa_i(t))$ was conducted on empirical rather than theoretical grounds, and created overfitting. Three mechanisms were used to guard against overfitting. Our earlier paper (electronically "published" at www.bschool.nus.edu.sg/depart/mk/bizcjk/fewa.htm.) began with four games (mixed strategies, continental divide, median action, and p-beauty contests). Frankly, several different specifications were explored to fit those four games best. Plausible functional forms were then 'frozen' and used on three more games (patent race, pot games, and traveler's dilemma). The fact that fEWA fits out-of-sample about as well as more complex theories in those three games (see Tables 2-3) suggests that searching across specifications on the first four games did not lead to much overfitting.

[31]This number is important because in some cases the payoff difference is large but the sample sizes which produce either model-consistent or model-inconsistent advice are small, so the statistical tests for differences lack power.

In addition, in public seminars and dissemination of our earlier draft (see its footnote 26) we encouraged other researchers to submit datasets to test the robustness of fEWA and other theories. We promised to run parallel analyses on all the models we initially estimated, and both report results and allow authors submitting data to report the results in their own papers. By the end of 2001, we received three such datasets.

The first dataset is from Kocher and Sutter (2000). Their data are choices made by individual and three-person groups in p-beauty contests. Competing groups appear to learn much faster than individuals. The second dataset is from Cabrales, Nagel, and Armenter (2001). They report data from a 2x2 coordination game with incomplete information, studying theories of "global games". In their games, players receive a private signal (knowing the distribution of private signals) which shows the payoffs in a signal-dependent stag-hunt coordination game. The game was designed so that iterated dominance leads subjects to select the risk-dominant outcome. The third dataset is from Sovik (2001). She reports data from a zero-sum betting game with asymmetric information. Two players simultaneously choose between a safe outside option and a zero-sum game based on their information sets. The information sets were chosen such that the game is dominance solvable. Players should figure out that they should never bet because they can only bet against players with better information. The "Groucho Marx Theorem" applies and the equilibrium is for no pair of agents to agree to bet in any information set. The observed rate of participation in the zero-sum game appears to be a function of the information set, which runs counter to the equilibrium prediction.

These three games check robustness of fEWA (and other models) in two directions: The Kocher and Sutter (2000) data stretch the models to explain choices made by *groups*; and the other two datasets are games of incomplete information. (All the seven games we studied in Section 3 are games with complete information.) Furthermore, since we announced a willingness to estimate models on *any* submitted data, there is no chance that we chose data which would bias results in any particular direction.

Table 7 reports results from the three new games. EWA fit best in LL for all three games and best in hit rate in two of the three games. fEWA did best in hit rate in the Kocher-Sutter beauty contests and was close to EWA in the remaining two. QRE was worst in Kocher and Sutter (2000) and Sovik (2001). Reinforcement was worst in the Cabrales et al global games. If fEWA was created by a specification search which overstated model performance in the first seven games, it should fit much worse in these three new games. In fact, it fits reasonably well. The functional forms of fEWA therefore show promise for predicting behavior in brand new games with incomplete information and games with choices made by groups rather than individuals.

Finally, fEWA seems to work almost as well as parametric EWA in modeling adaptive learning of short-term players in repeated trust and entrty games. In Camerer, Ho, and Chong (2002), we extend EWA to include sophisticaed players who believes others are learning. These sophisticated players who realize they are matched with others repeatedly often have incentives to "teach" as in the theory of repeated games. We show that our teaching model fits better than a generalized equilibrium model called quantal response equilibrium in explaing behaviors in repeated trust games. In Camerer, Ho, Chong and Weigelt (2002), we replace parametric EWA by fEWA (resulting in a saving of 15 learning parameters) and apply the teaching model to both the repeated trust and entry games (about 20,000 stage games). The teaching model with fEWA predicts behavior almost as well as the original teaching model. This provides an indirect but powerful robustnesss test for the fEWA model in two completely different games.

6 Conclusion

Learning is important for economics. Equilibrium theories are useful because they suggest a possible limit point of a learning process and permit comparative static analysis. But if learning is slow, or the time path of behavior selects one equilibrium out of many, a precise theory of equilibration is crucial for knowing when or which equilibrium will result.

In the last ten years, many theories of individual learning in games have been proposed and fit to data from laboratory games in which experimenters have good control over players' information and incentives. Some of these theories, particularly reinforcement learning and fictitious play or Cournot belief learning, are very simple (i.e., they have only one or two free parameters which have to be specified or estimated from data). Simple theories have the advantage of econometric parsimony but often miss important features of empirical data and, importantly, can be improved by adding features judiciously. Other theories are quite complex (e.g., Crawford, 1995; Stahl, 1999).

Because there are many theories, applied to different games using different scientific standards of proof or utility, some healthy controversy has emerged about which models are best for which purposes. Parametric EWA (EWA for short) takes a middle road by hybridizing features of reinforcement and belief learning, which necessarily adds parameters (Camerer and

Ho, 1999). Estimates across around 30 experimental data sets show that adding these features improves fit and predictive accuracy, but the parameter values which maximize fit are typically significantly different in different games. This finding raises the question of how to predict in advance which parameter values will fit best in a particular game.

The theory described in this paper, fEWA, replaces three parameters in the EWA learning models with three functions that change over time in response to experience. One function is a "motion detector" ϕ which goes up (limited by one) when behavior by other players is stable, and dips down (limited by zero) when there is surprising new behavior by others. When ϕ dips down, the effects of old experience (summarized in attractions which cumulate or average previous payoffs) is diminished by decaying the old attraction by a lot. The second function δ is simply ϕ divided by W, the minimal number of strategies in equilibrium. This function ties responsiveness to foregone payoffs to environmental stability, and also lowers δ in games with mixed equilibria ($W > 1$). The third function κ is an index of concentration of choices (a normalized Gini coefficient). This characterizes players who "explore" a lot (trying different strategies, yielding a low κ) and those who "exploit" by locking in to a single choice (high κ). fEWA is more parsimonious than most learning theories because it has only one free parameter– the response sensitivity λ.

We report fit and prediction of data from seven experimental games using fEWA, the parameterized EWA model, and three other models (belief learning, reinforcement with payoff variability, and quantal response equilibrium (QRE)).

Note that QRE and fEWA have one free parameter, reinforcement has two, belief learning has three, and EWA has five. We report both in-sample fit (penalizing more complex theories using the Bayesian information criterion) and out-of-sample predictive accuracy, to be sure that more complex models do not necessarily fit better.

There are three key results.

First, fEWA fits and predicts about as accurately as EWA in all seven games; and it produces functional parameter values for δ and ϕ which track the estimated values of fixed parameters quite closely across games. fEWA therefore represents one solution to the central problem of flexibly generating EWA-like parameters across games. Because fEWA generates sensible cross-game parameter variation automatically, it fits and predicts better than other models when games are pooled and common parameters are estimated.

Second, we propose a new criterion for judging the usefulness of theories, called economic value. A theory's economic value is the incremental profit a subject would earn from following the theory's advice rather than making their own choices. Most learning models add economic value in most games, although equilibrium theory does not. Either EWA or fEWA add the most economic value (and add positive value for a majority of subjects), in five or seven of the seven games, depending on how parameters are estimated to provide advice. Similar conclusions are drawn when boomerang and Lucas-critique effects are controlled by simply comparing profits from choices consistent with model advice with profits from inconsistent choices.

Third, the functions in fEWA are robust across games. The earlier draft of this paper used three games to validate the functional forms. This paper added three brand new games (after the first version was written and circulated) to test robustness. The basic conclusions are replicated in these games, which have incomplete information and choices are made by groups rather than individuals.

What's the bottom line? Because we used many criteria and games, it is not surprising that no one theory is always best by every criteria in every game. The results are sensitive to which games are used, but not particularly sensitive to performance criteria. In coordination and dominance-solvable games, either EWA or fEWA fit and predict best and add the most economic value. In mixed-strategy games reinforcement fits a little better than the EWA models by statistical criteria, and adds similar economic value. Belief models hardly ever fit best (fEWA is simpler and almost always fits better). And all learning theories fit reliably better than QRE.

A next step in this research is to apply fEWA to a wider set of games. We would also like to find some axiomatic underpinnings for the functions, which are admittedly ad hoc. Extending the ϕ function to exploit information about ordered strategies might prove useful. And since fEWA is so parsimonious, it is useful as a building block for extending learning theories to include sophistication (players anticipating that others are learning; see Stahl, 1999) and explain "teaching" behavior in repeated games (Camerer, Ho and Chong 2002; Cooper and Kagel, 2001).

The theory is developed to fit experimental data, but the bigger scientific payoff will come from application to naturally-occurring situations. If learning is slow, a precise theory of economic equilibration is just as useful for predicting what happens in the economy as a theory of equilibrium. For example, institutions for matching medical residents and medical schools, and analogous matching in college sororities and college bowl games, developed over decades (Roth and Xing, 1994). Bidders in eBay auctions learn to bid late to hide their information about an object's common value (Bajari and Hortacsu, 1999). Consumers learn over time what products they like (Ho and Chong, 2002). Learning in financial markets can generate excess volatility and returns predictability, which are otherwise anomalous in rational expectations models (Timmerman, 1993). Sargent (1999) argues that learning by policymakers about expectational Phillips' curves and the public's perceptions of inflation explains macroeconomic behavior in the last couple of decades. Good theories of learning should be able to explain these patterns and help predict how new institutions will evolve, how rapidly bidders learn to wait, and which new products will succeed. Applying fEWA, and other learning theories, to field domains is therefore an important goal of future research.

References

[1] Amaldoss, Wilfred and Ho, Teck-Hua, "EWA Learning in Games with Different Group Sizes," Marketing Department Working Paper, The Wharton School, 2001.

[2] Arthur, Brian, "Competing Technologies, Increasing Returns, and Lock-in by Historical Events," *Economic Journal*, 1989, 99, pp. 116-131.

[3] Arthur, Brian, "Designing Economic Agents That Act Like Human Agents: A Behavioral Approach to Bounded Rationality," *American Economic Review*, 1991, 81(2), pp. 353-359.

[4] Bajari, Patrick and Hortacsu, Ali, "Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from Ebay Auctions," Working Paper, Department of Economics, Stanford University, 1999.

[5] Basu, Kaushik, "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory," *American Economic Review*, May 1984, 84(2), pp. 391-395.

[6] Brown, George, "Iterative Solution of Games by Fictitious Play," in *Activity Analysis of Production and Allocation*, New York: John Wiley & Sons, 1951.

[7] Bush, Robert and Mosteller, Frederick, *Stochastic models for learning*, New York: Wiley, 1955.

[8] Cabrales, Antonio, Nagel, Rosemarie and Armenter, Roc, "Equilibrium Selection through Incomplete Information in Coordination Games: An Experimental Study," Universitat Pompeu Fabra working paper, 2001.

[9] Camerer, Colin F., *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton:Princeton University Press, 2002.

[10] Camerer, Colin and Ho, Teck-Hua, "Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity and Time-Variation," *Journal of Mathematical Psychology*, 1998, 42, pp. 305-326.

[11] Camerer, Colin and Ho, Teck-Hua, "Experience Weighted Attraction Learning in Normal Form Games," *Econometrica*, 1999, 67, pp. 827-873.

[12] Camerer, Colin and Ho, Teck-Hua, "Strategic Learning and Teaching in Games," in S. Hoch and H. Kunreuther, eds., *Wharton on Decision Making*, New York: Wiley, 2001.

[13] Camerer, Colin F., Teck-Hua Ho and Juin-Kuan Chong, "Sophisticated EWA Learning and Strategic Teaching in Repeated Games," *Journal of Economic Theory*, 2002, 104, 137-188.

[14] Camerer, Colin F., Teck-Hua Ho, Juin-Kuan Chong and Keith Weigelt, "Strategic Teaching and Equilibrium Models of Repeated Trust and Entry Games," CalTech Working Paper, 2002, http://www.hss.caltech.edu/ camerer/camerer.html

[15] Camerer, Colin, Hsia, David and Ho, Teck-Hua, "EWA Learning in Bilateral Call Markets," in *Experimental Business Research*, ed. by A. Rapoport and R. Zwick, in press.

[16] Capra, Monica, Goeree, Jacob, Gomez, Rosario and Holt, Charles, "Anomalous Behavior in a Traveler's Dilemma," *American Economic Review*, 89(3), June 1999, pp. 678-690.

[17] Chen, Yan and Khoroshilov, Yuri, "Learning under Limited Information," working paper, Department of Economics, University of Michigan, Ann Arbor, 2000.

[18] Cheung, Yin-Wong and Friedman, Daniel, "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, 1997, 19, pp. 46-76.

[19] Cooper, David and Kagel, John, "Learning and Transfer in Signalling Games," working paper, Department of Economics, Case Western Reserve University, 2001.

[20] Crawford, Vincent P., "Adaptive Dynamics in Coordination Games," *Econometrica*, 1995, 63, pp. 103-143.

[21] Crick, Francis, *What Mad Pursuit? A Personal View of Scientific Discovery*. New York: Basic Books, 1988.

[22] Cross, John, *A Theory of Adaptive Learning Economic Behavior*, New York: Cambridge University Press, 1983.

[23] Erev, Ido and Roth, Alvin E., "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 1998, 88(4), pp. 848-81.

[24] Erev, Ido, Bereby-Meyer, Yoella and Roth, Alvin E., "The Effect of Adding a Constant to All Payoffs: Experimental Investigation, and a Reinforcement Learning Model with Self-Adjusting Speed of Learning," *Journal of Economic Behavior and Organization*, 1999, 39, pp. 111-128.

[25] Fudenberg, Drew and Levine, David K., *The Theory of Learning in Games*, Boston: MIT Press, 1998.

[26] Gittins, John, *Multi-armed Bandit Allocation Indices*, New York: Wiley, 1989.

[27] Haruvy, Ernan and Stahl, Dale O., " An Empirical Model of Equilibrium Selection in Symmetric Normal-form Games," University of Texas Department of Economics Working Paper, January 1998.

[28] Harley, Calvin, "Learning the Evolutionary Stable Strategies," *Journal of Theoretical Biology*, 1981, 89, pp. 611-633.

[29] Heller, Dana and Sarin, Rajiv, "Parametric Adaptive Learning," University of Chicago Working Paper, 2000.

[30] Ho, Teck-Hua, Camerer, Colin and Chong, Juin-Kuan, "Economic Value of fEWA: A Functional Theory of Learning in Games," May, 2001. http://www.bschool.nus.edu.sg /depart/mk/bizcjk/fewa.htm

[31] Ho, Teck-Hua, Camerer, Colin and Weigelt, Keith, "Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests"," *American Economic Review*, 1998, 88, pp. 947-969.

[32] Ho, Teck-Hua and Chong, Juin-Kuan, "A Parsimonious Model of SKU Choice," *Journal of Marketing Research*, 2002, in press.

[33] Ho, Teck-Hua, Wang, Xin and Camerer, Colin, "Individual Differences in EWA Learning with Partial Payoff Information," Marketing Department Working Paper 99-010, The Wharton School, 1999.

[34] Hopkins, Edward, "Two Competing Models of how People Learn in Games," *Econometrica*, 2002, in press.

[35] Josephson, Jens, "A Numerical Analysis of the Evolutionary Stability of Learning Rules," Stockholm School of Economics SSE/EFI paper no. 474, 2001, http://swopec.hhs.se/hastef/abs/hastef0474.htm.

[36] Kahneman, Daniel, Knetsch, Jack L. and Thaler, Richard H.,"The Endowment Effect, Loss Aversion, and Status Quo Bias: Anomalies," *Journal of Economic Perspectives*, Winter 1991, 5(1), p.193-206.

[37] Kocher, Martin and Sutter, Matthias, "When the 'Decision Maker' Matters: Individual versus Team Behavior in Experimental 'Beauty-Contest' Games," University of Innsbruck working paper, 2000.

[38] LeDoux, Joseph, *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York : Simon & Schuster, 1996.

[39] McAllister, Patrick H., "Adaptive Approaches to Stochastic Programming," *Annals of Operations Research*, 1991, 30, pp. 45-62.

[40] McKelvey, Richard and Palfrey, Thomas, "Quantal Response Equilibria for Normal Form Games," Games and Economic Behavior, 1995, 10, 6-38.

[41] McKelvey, Richard, Palfrey, Thomas and Weber, Roberto, "The Effects of Payoff Magnitude and Heterogeneity on Behavior in 2x2 Games with Unique Mixed Strategy Equilibria," *Journal of Economic Behavior and Organization*, 2000, 42(4), pp.523-548.

[42] Mookerjhee, Dilip, and Sopher, Barry, "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 1994, 7, pp. 62-91.

[43] Mookerjhee, Dilip, and Sopher, Barry, "Learning and Decision Costs in Experimental Constant-sum Games," *Games and Economic Behavior*, 1997, 19, pp. 97-132.

[44] Myung, In Jae, "The importance of complexity inn model selection," *Journal of Mathematical Psychology*, 2000, 44, pp. 190-204.

[45] Pratt, John, Wise, David and Zeckhauser, Richard, "Price Differences in Almost Competitive Markets," *Quarterly Journal of Economics,*, May 1979, 93(2), pp. 189-211.

[46] Rapoport, Amnon and Erev, Ido, “Coordination, “magic”, and reinforcement learning in a market entry game,” *Games and Economic Behavior*, 1998, 23, pp. 146-175.

[47] Rapoport, Amnon and Amaldoss, Wilfred, “Mixed Strategies and Iterative Elimination of Strongly Dominated Strategies: An Experimental Investigation of States of Knowledge,” *Journal of Economic Behavior and Organization*, 2000, 42, pp. 483-521.

[48] Robinson, Julia, “An Iterative Method of Solving a Game,” *Annals of Mathematics*, 1951, 54(2), pp. 296-301.

[49] Roth, Alvin E. and Xing, Xiaolin, “Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions,” *American Economic Review*, 84, September, 1994, 992-1044.

[50] Roth, Alvin E. and Erev, Ido, “Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term,” *Games and Economic Behavior*, 8(1), 1995, pp. 164-212.

[51] Roth, Alvin, Barron, Greg, Erev, Ido and Slonim, Robert, “Equilibrium and Learning in Economic Environments: the Predictive Value of Approximations,” Harvard University Working Paper, 2002.

[52] Salmon, Tim, “Evidence for Learning to Learn Behavior in Normal Form Games,” Caltech Working Paper, 1999.

[53] Samuelson, William and Zeckhauser, Richard, ”Status Quo Bias in Decision Making,” *Journal of Risk and Uncertainty*, March 1988, 1, pp. 7-59.

[54] Sargent, Thomas, *The Conquest of American Inflation*, Princeton: Princeton University Press, 1999.

[55] Selten, Reinhard and Stoecker, Rolf, “End Behavior in Sequences of Finite Prisoner’s Dilemma Supergames: A Learning Theory Approach,” *Journal of Economic Behavior and Organization*, 1986, 7, pp. 47-70.

[56] Shachat, Jason, ”Mixed Strategy Play and the Minimax Hypothesis,” *Journal of Economic Theory*, 2002, 104, pp. 189-226.

[57] Sovik, Ylva, “Impossible Bets: An Experimental Study,” University of Oslo working paper, 2001.

[58] Stahl, Dale O, "Boundedly Rational Rule Learning in a Guessing Game," *Games and Economic Behavior*, 1996, 16, pp. 303-330.

[59] Stahl, Dale O., "Sophisticated Learning and Learning Sophistication," University of Texas at Austin Working Paper, 1999.

[60] Stahl, Dale O, "Local Rule Learning: Theory and Evidence," *Games and Economic Behavior*, forthcoming.

[61] Sutton, Richard and Barto, Andrew, *Reinforcement Learning: An Introduction*, Boston: MIT Press, 1998.

[62] Timmerman, Allan G, "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices.' *Quarterly Journal of Economics*, November 1993, 108, pp. 1135-1145.

[63] Van Huyck, John, Battalio, Raymond and Beil, Richard, "Tacit Cooperation Games, Strategic Uncertainty, and Coordination Failure," *The American Economic Review*, 1990, 80, pp. 234-248.

[64] Van Huyck, John, Battalio, Raymond and Rankin, Frederick, "Selection Dynamics and Adaptive Behavior Without Much Information," Working Paper, Department of Economics, Texas A & M University, 1996.

[65] Van Huyck, John, Cook, Joseph and Battalio, Raymond, "Adaptive Behavior and Coordination Failure," *Journal of Economic Behavior and Organization*, 1997, 32, pp. 483-503.

7 Appendix

7.1 Calculating the Initial Attractions

We "burn in" the initial attractions $A^j(0), \forall j$, by using the actual observed frequency of choices by all subjects in the first period. The same initial attractions are used for all subjects, except for games with more than 2 players such as continental divide, median action and p-beauty contest. For the exceptions, each group of players has a different initial attractions based on the group's observed frequencies. For a particular response sensitivity λ, the initial attractions are chosen so that the predicted probabilities of choices match the actual relative frequencies of choices.

Denote the empirically observed frequency of strategy j in the first period by f^j. Then initial attractions are recovered from the equations

$$\frac{e^{\lambda \cdot A^j(0)}}{\sum_k e^{\lambda \cdot A^k(0)}} = f^j, j = 1, \ldots, m. \tag{7.1}$$

(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

$$A^j(0) - A^k(0) = \frac{1}{\lambda} ln(f^j) - \frac{1}{\lambda} ln(f^k), \quad j, k = 1, \ldots, m \tag{7.2}$$

We fix the initial attraction of the strategy j with the lowest frequency $A^j(0)$ to a constant value for identification. Frequently, the lowest frequency is zero. We circumvent this problem by adding a constant $\frac{W}{m}$ to all frequencies and renormalizing them.

$$\tilde{f^j} = \frac{f^j + \frac{W}{m}}{\sum_k f^k + \frac{W}{m}}, \quad j = 1, \ldots, m \tag{7.3}$$

$\frac{W}{m}$ is chosen to reflect the relative proportion of equilibrium points with respect to number of strategies. With $\tilde{f^j}$ in place of f^j in (7.2), we then solve for the other attractions as a function of λ and the modified frequencies $\tilde{f^j}$.

To ensure no model obtains any unfair advantage from the burn in procedure, we use (7.1) as the first period prediction for all models.

Since the model specifications for fEWA, EWA and Belief-based learning have been discussed in the text, we only provide the model specifications for the remaining models, Quantal Response and Reinforcement with Payoff Variability, below.

7.2 Quantal Response Model

The updating rule and predicted probability are given as follows:

$$A_i^j(t) = \sum_{k=1}^{m_{-i}} P_{-i}^k(t+1) \cdot \pi_i(s_i^j, s_{-i}^k(t)) \tag{7.4}$$

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}} \tag{7.5}$$

As evident from (7.5), the predicted probability is a function of other player(s) predicted probabilities. For a given sensitivity parameter λ, predicted probabilities are derived from solving N nonlinear simultaneous equations. We solve the nonlinear simultaneous equations numerically by iterative substitutions until we converge to a set of consistent predicted probabilities.

7.3 Reinforcement Model with Payoff Variability

The updating rule is:

$$A_i^j(t) = \frac{(N(0) + C_{ij}(t) - 1) \cdot A_i^j(t-1) + I(s_i^j, s_i(t)) \cdot \pi_i(s_i^j, s_{-i}(t))}{N(0) + C_{ij}(t)} \tag{7.6}$$

where $C_{ij}(t)$ (with $C_{ij}(0) = 0, \ \forall i, j$) is updated as follows:

$$C_{ij}(t) = \begin{cases} C_{ij}(t-1) + 1 & \text{if } j \text{ is chosen in } t \\ C_{ij}(t-1) & \text{if } j \text{ is not chosen in } t \end{cases} \tag{7.7}$$

In addition, λ is replaced by $\frac{\lambda}{S_i(t)}$ where

$$S_i(t) = \frac{(t - 1 + m \cdot N(0))S_i(t-1) + |\overline{\pi_i(t-1)} - \pi_i(s_i(t), s_{-i}(t))|}{t + m \cdot N(0)} \tag{7.8}$$

where m is the number of strategies and $A_i(t)$ is updated as follows:

$$\overline{\pi_i(t)} = \frac{(t - 1 + m \cdot N(0))\overline{\pi_i(t-1)} + \pi_i(s_i(t), s_{-i}(t))}{t + m \cdot N(0)} \tag{7.9}$$

where $\overline{\pi_i(0)}$ is the expected payoff given random choice.

Instead of assuming random choices by other players in the computation of $\overline{\pi_i(0)}$, we use empirical distribution of other players' first period choices to increase the potency of the model. This also ensures that the Reinforcement Model with Payoff Variability is placed on the same footing as other models where first period is used to burn in initial attractions. $S_i(0)$ is the expected absolute difference between payoff from each strategy and $\overline{\pi_i(0)}$.

7.4 Parameter Estimates and Functional Values

Table A.1 gives the parameter estimates and their standard errors of all learning models. Note that the standard errors are small, suggesting that these parameters are statistically different from zero. Table A.2 shows the inter-quartile ranges of fEWA functional values across time and subjects. Except for κ in some games, the ranges are relatively small. Table A.3 gives the parameter estimates for the three new games that are used to test the robustness of fEWA.

Table 1: A Description of the Seven Games Used in the Estimation of Various Learning Models

Game	Number of Players	Number of Strategies	Number of Pure Strategy Equilibria	Number of Subjects	Number of Rounds	Matching Protocol	Experimental Treatment	Description of Games
Mixed Strategies Mookerjhee and Sopher (1997)	2	4,6	0	80	40	Fixed	Stake Size	A constant-sum game with unique mixed strategy equilibrium.
Patent Race Rapoport and Amaldoss (2000)	2	5,6	0	36	80	Random	Strong vs Weak	Strong (weak) player invests between 0 and 5 (0 and 4) and the higher investment wins a fixed prize.
Continental Divide Van Huyck et al. (1997)	7	14	2	70	15	Fixed	None	A coordination game with two pure strategy equilibria
Median Action Van Huyck et al. (1990)	9	7	7	54	10	Fixed	None	A order-statistic game with individual payoff decreases in the distance between individual choice and the median
Pot Games Amaldoss and Ho (2001)	3,6,9,18	2	1	84	25 (manual) 28 (computer)	Fixed	Number of Players	An entry game where players must decide which of the two ponds of sizes 2n and n they wish to enter. Payoff is the ratio of the pond size and number of entries.
p-Beauty Contest Ho et al. (1998)	7	101	1	196	10	Fixed	Experienced vs. Inexperienced	Players simultaneously choose a number from 0 to 100 and the winner whose number is closet to p (<1) times the group average
Traveler's Dilemma Capra et al. (1999)	2	121[1]	1	52	10	Random	Penalty Size	Players choose claims between 80 and 200. Both players get lower claim but the high-claim player pays a penalty to the low-claim player.

Note 1: Continuous strategies of 80 to 200 are discretized to 121 integer strategies

Table 2: Model Fit (% Hit Rate, BIC and Log Likelihood)

In-sample Calibration

	Sample	Random	fEWA		EWA		Belief-based		Reinforcement with PV		QRE	
	Size[5]	%Hit	%Hit[1,4]	BIC[2]	%Hit	BIC	%Hit	BIC	%Hit	BIC	%Hit	BIC
Mixed Strategies	2240	21%	40%	-3192	**41%**	-3074	38%	-3129	*41%*	**-3051**	30%	-3342
Patent Race [6]	4000	18%	*62%*	-4442	*61%*	-4411	52%	-5506	**62%**	**-4367**	38%	-6682
Continental Divide	735	7%	*50%*	-1081	**51%**	**-1062**	30%	-1288	47%	-1293	7%	-1890
Median Action	380	14%	69%	-313	**75%**	**-272**	*74%*	-348	69%	-343	50%	-560
Pot Games	1478	50%	**68%**	**-905**	*67%*	-937	65%	-982	*66%*	-907	62%	-1018
p-Beauty Contest	1380	1%	**13%**	-4567	*13%*	**-4544**	*13%*	-4571	10%	-5741	3%	-5849
Traveler's Dilemma	360	1%	**51%**	-890	*50%*	**-873**	31%	-1120	46%	-1069	32%	-1613
Pooled [3]	10573	20%	**51%**	**-15389**	48%	-15906	40%	-17960	44%	-20182	33%	-21055

Out-of-sample Validation

	Sample	Random	fEWA		EWA		Belief-based		Reinforcement with PV		QRE	
	Size	%Hit	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL
Mixed Strategies	960	21%	**36%**	**-1382**	*36%*	-1387	*34%*	-1404	*33%*	-1392	*35%*	-1398
Patent Race	1760	18%	*65%*	-1897	**65%**	-1878	52%	-2279	*65%*	**-1864**	19%	-2978
Continental Divide	315	7%	*47%*	-470	**47%**	**-460**	25%	-564	44%	-573	5%	-808
Median Action	160	14%	74%	-104	*79%*	**-83**	**82%**	-95	74%	-105	49%	-196
Pot Games	739	50%	**70%**	-436	**70%**	-437	66%	-471	*70%*	**-432**	65%	-505
p-Beauty Contest	580	1%	**8%**	-2119	*6%*	**-2046**	*7%*	-2037	*7%*	-2498	3%	-2507
Traveler's Dilemma	160	1%	**46%**	-446	*43%*	**-445**	36%	-465	*41%*	-564	31%	-699
Pooled	4674	20%	**51%**	**-6853**	49%	-7100	40%	-7935	46%	-9128	36%	-9037

Note 1: Number of hits counts the occasions when prob(chosen strategy) = maximum (predicted probabilities). Each count is adjusted by number of strategies sharing the maximum.

Note 2: BIC (Bayesian Information Criterion) is given by LL - (k/2)*log(N*T) where k is the number of parameters, N is the number of subjects and T is the number of periods.

Note 3: A common set of parameters, except game-specific lambda, is estimated for all games. Each game is given equal weight in LL estimation.

Note 4: Entries in **bold** denote the best measures for each game. In case of hit rate, multiple models might share the top rank when differences in hit rates of these models are not statistically significant by McNemar test (Chi-sq at 5%); these entries are *underlined*.

Note 5: Calibrated on all observations for 70% of the subjects instead of 70% observations of all subjects as in Camerer and Ho (1999).

Note 6: Games in **bold** were added after the functional forms of the parameters were adopted.

Table 3: Model Robustness (Out-of-sample Prediction for Entire Game)

Out-of-sample Prediction using Out-of-game Estimates[1]

	Sample	fEWA		EWA		Belief-based		Reinforcement with PV		QRE	
	Size	%Hit [2]	LL	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL
Mixed Strategies	3200	**39%**	**-4662**	34%	-4867	35%	-4832	*38%*	-4697	31%	-5049
Patent Race	5760	*62%*	-8009	59%	-9296	55%	-9112	**63%**	**-6588**	39%	-9745
Continental Divide	1050	48%	-1741	**50%**	**-1635**	27%	-2147	32%	-2403	6%	-2695
Median Action	540	69%	-523	**74%**	-491	60%	-711	70%	**-479**	50%	-990
Pot Games	2217	**68%**	-3976	67%	-5084	65%	-3474	67%	**-1387**	63%	-1491
p-Beauty Contest	1960	**12%**	**-6681**	10%	-6819	8%	-7715	9%	-11361	3%	-8342
Traveler's Dilemma	520	**50%**	**-1825**	46%	-1874	33%	-1933	36%	-1841	21%	-2325

Note 1:Prediction for a game is made using out-of-game estimates derived from pooling the other 6 games.

Note 2: Entries in **bold** denote the best measures for each game. In case of hit rate, multiple models might share the top rank when differences in hit rates of these models are not statistically significant by McNemar test (Chi-sq at 5%); these entries are *underlined*.

Table 4: Payoffs in ‘continental divide’ experiment, Van Huyck et al. (1997)

	Median Choice													
choice	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	45	49	52	55	56	55	46	-59	-88	-105	-117	-127	-135	-142
2	**48**	53	58	62	65	66	61	-27	-52	-67	-77	-86	-92	-98
3	**48**	**54**	***60***	**66**	70	74	72	1	-20	-32	-41	-48	-53	-58
4	43	51	58	65	**71**	**77**	80	26	8	-2	-9	-14	-19	-22
5	35	44	52	60	69	**77**	**83**	46	32	25	19	15	12	10
6	23	33	42	52	62	72	82	62	53	47	43	41	39	38
7	7	18	28	40	51	64	78	75	69	66	64	63	62	62
8	-13	-1	11	23	37	51	69	83	81	80	80	80	81	82
9	-37	-24	-11	3	18	35	57	88	89	91	92	94	96	98
10	-65	-51	-37	-21	-4	15	40	**89**	**94**	98	101	104	107	110
11	-97	-82	-66	-49	-31	-9	20	85	**94**	**100**	105	110	114	119
12	-133	-117	-100	-82	-61	-37	-5	78	91	99	**106**	***112***	**118**	**123**
13	-173	-156	-137	-118	-96	-69	-33	67	83	94	103	110	117	**123**
14	-217	-198	-179	-158	-134	-105	-65	52	72	85	95	104	112	120

Table 5: Economic Values with In-game and Out-of-game Estimates

Total Payoff and Percentage Improvement for Bionic Subjects using Out-of-game Estimates[1]

	Observed Payoff	Ex-post Maximum %Improve	fEWA %Improve [3]	EWA %Improve	Belief-based %Improve	Reinforcement with PV %Improve	QRE %Improve
Mixed Strategies	334	100.0%	**7.5%**	3.0%	1.1%	5.8%	-1.8%
Patent Race	467	44.2%	1.7%	1.2%	1.3%	**2.9%**	1.2%
Continental Divide[2]	837	6.6%	5.0%	**5.2%**	4.5%	-9.4%	-30.5%
Median Action[2]	503	1.8%	1.5%	**1.5%**	1.2%	1.3%	-1.0%
Pot Games	4244	29.9%	-2.7%	-1.1%	-1.3%	-1.9%	**9.9%**
p-Beauty Contest[2]	519	585.4%	**49.9%**	40.8%	26.7%	-7.2%	-64.0%
Traveler's Dilemma	541	26.2%	**10.3%**	9.8%	9.4%	3.4%	2.7%

Total Payoff and Percentage Improvement for Bionic Subjects using In-game Estimates[1]

	Observed Payoff	Ex-post Maximum %Improve	fEWA %Improve	EWA %Improve	Belief-based %Improve	Reinforcement with PV %Improve	QRE %Improve
Mixed Strategies	334	100.0%	13.0%	**15.9%**	14.0%	13.0%	-1.8%
Patent Race	467	44.2%	**2.4%**	2.0%	2.4%	2.0%	-9.6%
Continental Divide[2]	837	6.6%	**4.8%**	4.9%	4.9%	3.2%	-32.0%
Median Action[2]	503	1.8%	**1.5%**	**1.5%**	**1.4%**	1.5%	-1.0%
Pot Games	4244	29.9%	7.7%	**11.1%**	7.4%	9.3%	9.9%
p-Beauty Contest[2]	519	585.4%	48.8%	**49.1%**	51.1%	-57.5%	-65.0%
Traveler's Dilemma	541	26.2%	11.1%	**11.1%**	9.5%	8.2%	7.9%

Note 1: We assume that each "bionic" subject uses the respective model to predict other's behavior and best responds with the strategy that yields the highest expected payoff.

Note 2: The expected value of each strategy in these games is computed with 1000 simulated instances for a given round due to the combinatorial nature of determining payoff and probability of each strategy

Note 3: Entries in **bold** denote the best improvement for each game.

Table 6: Economic Values (Expected Payoff Differences of Model-consistent Choices vs Other Choices) with In-game Parameter Estimates

	fEWA			EWA			Belief-based			Reinforcement with PV			QRE		
	Π_{diff} [1]	*p*-value [3]	%BR [4]	Π_{diff}	*p*-value	%BR	Π_{diff}	*p*-value	%BR	Π_{diff}	*p*-value	%BR	Π_{diff}	*p*-value	%BR
Mixed Strategies	0.02	0.00	32%	0.02	0.00	31%	0.02	0.00	27%	0.02	0.00	32%	0.00	0.38	31%
Patent Race	0.01	0.00	37%	0.01	0.00	37%	0.02	0.00	42%	0.01	0.00	37%	0.01	0.00	34%
Continental Divide[2]	0.14	0.00	37%	0.15	0.00	37%	0.15	0.00	36%	0.10	0.00	38%	-0.21	1.00	5%
Median Action[2]	0.05	0.00	81%	0.05	0.00	81%	0.05	0.00	81%	0.05	0.00	81%	-0.06	1.00	51%
Pot Games	0.25	0.00	63%	0.23	0.00	60%	0.12	0.00	59%	0.35	0.00	64%	0.40	0.00	63%
p-Beauty Contest[2]	0.19	0.03	4%	0.07	0.20	5%	0.01	0.43	5%	-0.07	0.68	1%	-0.17	0.93	1%
Traveler's Dilemma	0.44	0.00	26%	0.43	0.00	26%	0.44	0.00	25%	0.41	0.00	25%	0.10	0.00	32%

Note 1: We split the choices into 2 sets: one set contains choices that best respond to predicted behavior of others while the other set contains choices that do not. We derive the difference in mean payoff between the 2 sets.

Note 2: The expected value of each strategy in these games is computed with 1000 simulated instances for a given round because of the combinatorial nature of detrmining payoff and probability of the various strategy combinations.

Note 3: We assume unequal variances for the 2 sets of choices. Student-t distribution is used to derive the p-value.

Note 4: Percentage of choices that best respond to model-predicted behaviors of others.

Table 7: Model Robustness Test for 3 New Data Sets (% Hit Rate, BIC and Log Likelihood)

In-sample Calibration

	Sample Size[3]	fEWA		EWA		Belief-based		Reinforcement with PV		QRE	
		%Hit[1,4]	BIC[2]	%Hit	BIC	%Hit	BIC	%Hit	BIC	%Hit	BIC
Cabrales, Nagel & Armenter (2001)	1200	86%	-364	**88%**	**-353**	85%	-407	83%	-423	86%	-453
Kocher & Sutter (2000) [individual]	100	**8%**	-400	7%	-365	7%	**-362**	5%	-445	1%	-459
Kocher & Sutter (2000) [group]	100	**4%**	-403	3%	-343	3%	**-339**	2%	-437	1%	-455
Sovik (2001)	864	65%	-568	**66%**	**-553**	64%	-552	64%	-554	58%	-579

Out-of-sample Validation

	Sample Size	fEWA		EWA		Belief-based		Reinforcement with PV		QRE	
		%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL
Cabrales, Nagel & Armenter (2001)	400	87%	-124	**88%**	**-112**	88%	-126	87%	-130	89%	-124
Kocher & Sutter (2000) [individual]	40	**5%**	-150	3%	**-136**	3%	-137	3%	-170	1%	-180
Kocher & Sutter (2000) [group]	40	**8%**	-159	3%	**-142**	3%	**-142**	5%	-182	1%	-180
Sovik (2001)	288	70%	-174	**78%**	**-162**	74%	-165	73%	-170	71%	-175

Total Log-Likelihood (in and out-of-sample) and Hit Rate

	Sample Size[3]	fEWA		EWA		Belief-based		Reinforcement with PV		QRE	
		%Hit[1]	LL	%Hit	LL	%Hit	LL	%Hit	LL	%Hit	LL
Cabrales, Nagel & Armenter (2001)	1600	86%	-484	**88%**	**-448**	86%	-523	84%	-546	87%	-574
Kocher & Sutter (2000) [individual]	140	**8%**	-550	6%	**-501**	6%	-498	4%	-615	1%	-639
Kocher & Sutter (2000) [group]	140	**5%**	-562	3%	-485	3%	**-480**	3%	-619	1%	-635
Sovik (2001)	1152	66%	-738	**69%**	**-699**	67%	-707	66%	-717	62%	-750

Note 1: Number of hits counts the occasions when prob(chosen strategy) = maximum (predicted probabilities). Each count is adjusted by number of strategies sharing the maximum.

Note 2: BIC (Bayesian Information Criterion) is given by LL - (k/2)*log(N*T) where k is the number of parameters, N is the number of subjects and T is the number of periods.

Note 3: Calibrated on 70% of the subjects and validated on remaining 30%.

Note 4: Entries in **bold** denote the best measures for each game.

Table A.1: Parameter Estimates of Learning Models

fEWA [1]

	φ		κ		δ		N0 [3]		λ [5]	
Mixed Strategies	0.89	-	0.52	-	0.28	-	1.00	-	3.78	0.17
Patent Race	0.89	-	0.72	-	0.32	-	1.00	-	7.87	0.17
Continental Divide	0.69	-	0.77	-	0.69	-	1.00	-	4.46	0.18
Median Action	0.85	-	0.78	-	0.85	-	1.00	-	5.00	0.33
Pot Games	0.80	-	0.44	-	0.44	-	1.00	-	0.33	0.03
p-Beauty Contest	0.58	-	0.82	-	0.58	-	1.00	-	2.11	0.05
Traveler's Dilemma	0.63	-	0.84	-	0.63	-	1.00	-	4.99	0.20
Pooled	0.76	-	0.64	-	0.48	-	1.00	-	3.26	0.05

EWA

	φ		κ		δ		N0 [4]		λ	
Mixed Strategies	0.98	0.00	1.00	0.04	0.27	0.07	0.82	0.00	1.15	0.10
Patent Race	0.92	0.01	0.05	0.02	0.36	0.25	1.37	0.02	42.21	4.57
Continental Divide	0.74	0.03	1.00	0.02	0.73	0.09	0.25	0.00	3.98	0.30
Median Action	0.71	0.07	1.00	0.02	0.89	0.00	0.00	0.00	8.90	1.12
Pot Games	0.81	0.04	1.00	0.09	0.42	0.00	0.00	0.00	0.19	0.03
p-Beauty Contest	0.36	0.02	0.00	0.04	0.78	0.05	1.56	0.00	3.44	0.00
Traveler's Dilemma	0.77	0.02	1.00	0.02	0.53	0.07	0.62	0.00	3.53	0.22
Pooled [2]	0.78	0.01	0.99	0.01	0.49	0.00	0.01	0.00	2.95	0.02

Belief-based

	φ		κ [3]		δ [3]		N0 [4]		λ	
Mixed Strategies	1.00	0.00	0.00	-	1.00	-	70.58	2.12	43.40	0.24
Patent Race	1.00	0.00	0.00	-	1.00	-	27.77	141.07	85.09	0.01
Continental Divide	0.99	0.05	0.00	-	1.00	-	1.05	0.25	14.74	1.26
Median Action	1.00	0.00	0.00	-	1.00	-	5.86	9.10	75.84	0.08
Pot Games	0.98	0.04	0.00	-	1.00	-	0.40	0.17	0.87	0.13
p-Beauty Contest	0.33	0.02	0.00	-	1.00	-	0.85	0.36	2.57	0.08
Traveler's Dilemma	0.85	0.01	0.00	-	1.00	-	6.69	1.50	13.97	0.85
Pooled	0.81	0.02	0.00	-	1.00	-	5.36	0.49	11.59	0.48

Reinforcement with PV

	φ [3]		κ [3]		δ [3]		N0		λ	
Mixed Strategies	1.00	-	0.00	-	0.00	-	31.47	0.91	4.39	0.52
Patent Race	1.00	-	0.00	-	0.00	-	6.58	0.03	1.48	0.05
Continental Divide	1.00	-	0.00	-	0.00	-	1.81	0.13	2.59	0.03
Median Action	1.00	-	0.00	-	0.00	-	3.95	0.24	1.27	0.19
Pot Games	1.00	-	0.00	-	0.00	-	0.49	0.03	0.39	0.03
p-Beauty Contest	1.00	-	0.00	-	0.00	-	1.78	0.01	0.22	0.01
Traveler's Dilemma	1.00	-	0.00	-	0.00	-	3.90	0.03	3.21	0.03
Pooled	1.00	-	0.00	-	0.00	-	188.10	0.24	24.49	0.10

QRE

	φ [3]		κ [3]		δ [3]		N0 [3]		λ	
Mixed Strategies	0.00	-	0.00	-	0.00	-	0.00	-	10.53	0.08
Patent Race	0.00	-	0.00	-	0.00	-	0.00	-	4.67	0.04
Continental Divide	0.00	-	0.00	-	0.00	-	0.00	-	0.50	0.01
Median Action	0.00	-	0.00	-	0.00	-	0.00	-	20.00	0.00
Pot Games	0.00	-	0.00	-	0.00	-	0.00	-	0.04	0.00
p-Beauty Contest	0.00	-	0.00	-	0.00	-	0.00	-	0.00	0.00
Traveler's Dilemma	0.00	-	0.00	-	0.00	-	0.00	-	20.00	0.00
Pooled	0.00	-	0.00	-	0.00	-	0.00	-	7.78	0.01

Note 1: Average parameters across subjects and time.

Note 2: For all models except fEWA, a common set of estimates, except lambdas, is estimated for all games pooled.

Note 3: Fixed parameters

Note 4: N0 bounded by 1/(1-phi(1-kappa)).

Note 5: Payoffs in all games have been rescaled to USD equivalent. Average of game specific lambda is reported for pooled games.

Table A.2: Variations of fEWA Functional Values

Interquartile Range Across Time[1]

	ϕ		κ		δ	
	Median[3]	Range[4]	Median	Range	Median	Range
Mixed Strategies	0.91	0.0396	0.50	0.1027	0.30	0.1137
Patent Race	0.90	0.0200	0.73	0.0151	0.31	0.0078
Continental Divide	0.69	0.0806	0.78	0.0165	0.69	0.0806
Median Action	0.91	0.1875	0.81	0.1513	0.91	0.1875
Pot Games	0.81	0.0635	0.44	0.1329	0.44	0.0464
p-Beauty Contest	0.58	0.0375	0.88	0.0673	0.58	0.0375
Traveler's Dilemma	0.62	0.0722	0.89	0.0926	0.62	0.0722
Pooled	0.87	0.1218	0.57	0.2610	0.35	0.1512

Interquartile Range Across Subjects[2]

	ϕ		κ		δ	
	Median	Range	Median	Range	Median	Range
Mixed Strategies	0.89	0.0357	0.52	0.1504	0.27	0.1026
Patent Race	0.89	0.0133	0.74	0.2036	0.32	0.0044
Continental Divide	0.68	0.0820	0.78	0.0835	0.68	0.0820
Median Action	0.85	0.0000	0.79	0.0432	0.85	0.0000
Pot Games	0.78	0.0167	0.38	0.3247	0.43	0.0074
p-Beauty Contest	0.58	0.0156	0.82	0.0139	0.58	0.0156
Traveler's Dilemma	0.62	0.0830	0.84	0.0121	0.62	0.0830
Pooled	0.76	0.2694	0.82	0.2057	0.57	0.2664

Note 1: Average value for each time period is calculated, then the 25% and 75% percentile are used to derive the range.
Note 2: Average value for each subject is calculated, then the 25% and 75% percentile are used to derive the range.
Note 3: Overall medians on the average values calculated with respect to Note 1 and 2.
Note 4: The 25% and 75% percentile are respectively 0.5*Range below and above the median.

Table A.3: Parameter Estimates of Learning Models for 3 New Data Sets[1]

		f		k		d		N0[3]		l	
fEWA [2]	Cabrales, Nagel & Armenter (2001)	0.86	-	0.58	-	0.86	-	1.00	-	9.95	0.05
	Kocher & Sutter (2000) [individual]	0.58	-	0.73	-	0.58	-	1.00	-	0.49	0.06
	Kocher & Sutter (2000) [group]	0.58	-	0.73	-	0.58	-	1.00	-	0.43	0.05
	Sovik (2001)	0.74	-	0.39	-	0.74	-	1.00	-	2.38	0.17
EWA	Cabrales, Nagel & Armenter (2001)	0.63	0.02	0.99	0.04	0.27	0.00	0.01	0.00	8.19	0.00
	Kocher & Sutter (2000) [individual]	0.71	0.06	0.00	0.00	0.88	0.02	0.32	0.00	0.59	0.03
	Kocher & Sutter (2000) [group]	0.64	0.09	0.09	7.86	1.00	48.43	0.44	1.08	0.53	0.05
	Sovik (2001)	1.04	0.00	0.13	0.05	0.59	0.42	0.88	0.00	5.26	0.10
Belief-based	Cabrales, Nagel & Armenter (2001)	0.83	0.02	0.00	-	1.00	-	3.22	1.27	28.89	0.08
	Kocher & Sutter (2000) [individual]	0.70	0.05	0.00	-	1.00	-	0.31	0.05	1.06	0.05
	Kocher & Sutter (2000) [group]	0.64	0.08	0.00	-	1.00	-	0.44	0.91	0.52	0.08
	Sovik (2001)	1.00	0.00	0.00	-	1.00	-	1.89	0.06	8.13	0.24
Reinforcement with PV	Cabrales, Nagel & Armenter (2001)	1.00	-	0.00	-	0.00	-	316.74	0.01	119.57	0.01
	Kocher & Sutter (2000) [individual]	1.00	-	0.00	-	0.00	-	0.99	0.16	0.15	0.41
	Kocher & Sutter (2000) [group]	1.00	-	0.00	-	0.00	-	0.89	0.12	0.19	0.15
	Sovik (2001)	1.00	-	0.00	-	0.00	-	1.89	0.91	2.09	0.52
QRE	Cabrales, Nagel & Armenter (2001)	0.00	-	0.00	-	0.00	-	0.00	-	11.37	3.18
	Kocher & Sutter (2000) [individual]	0.00	-	0.00	-	0.00	-	0.00	-	0.50	0.00
	Kocher & Sutter (2000) [group]	0.00	-	0.00	-	0.00	-	0.00	-	0.00	0.01
	Sovik (2001)	0.00	-	0.00	-	0.00	-	0.00	-	2.75	0.04

Note 1: Underscored are fixed parameters

Note 2: Average parameters across subjects and time.

Note 3: N0 is bounded by 1/(1-phi(1-kappa)) for EWA and Belief-based Models

Figure 1: EWA's Model Parametric Space

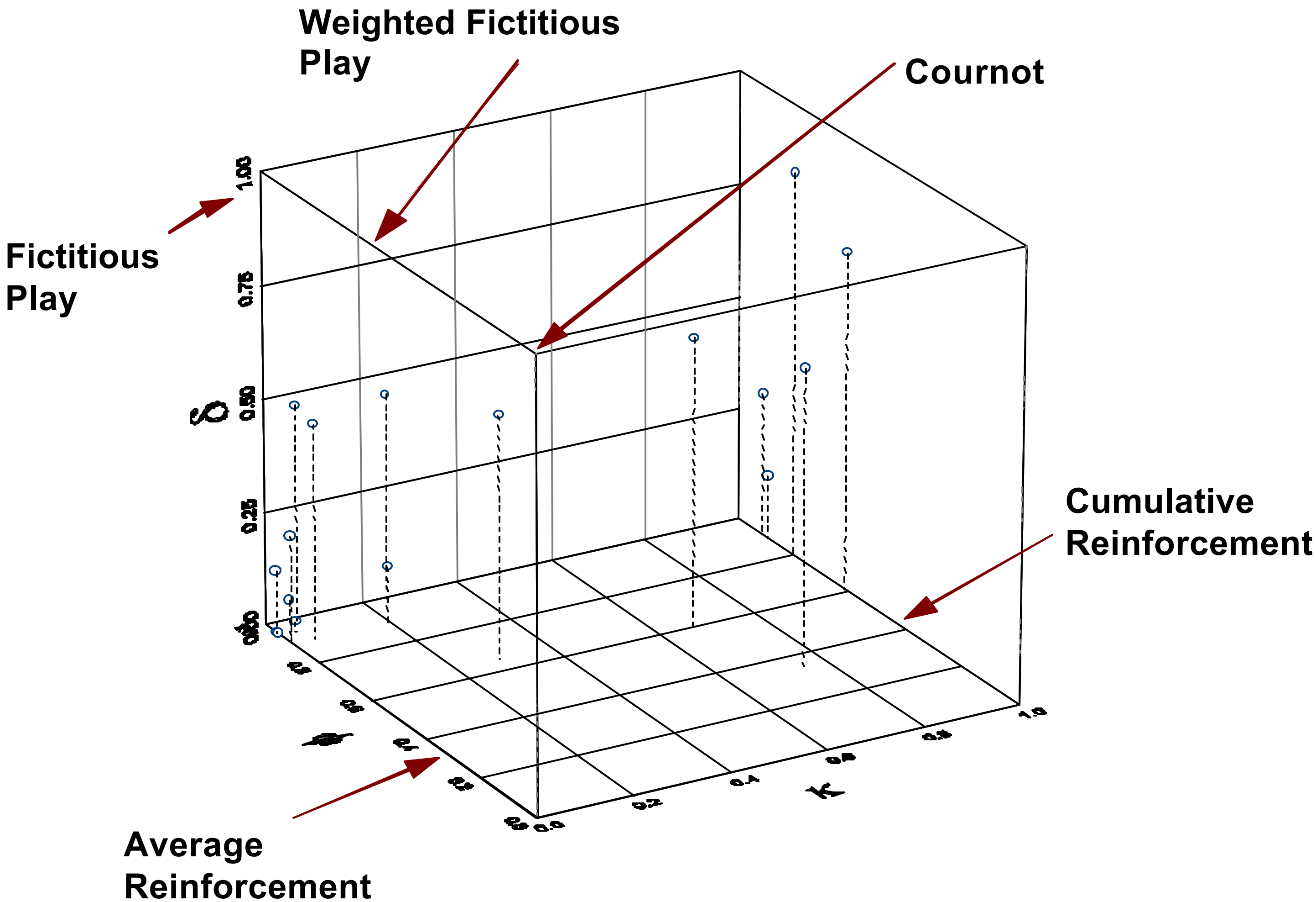

Figure 2: Parametric Plot of fEWA vs Parametric EWA

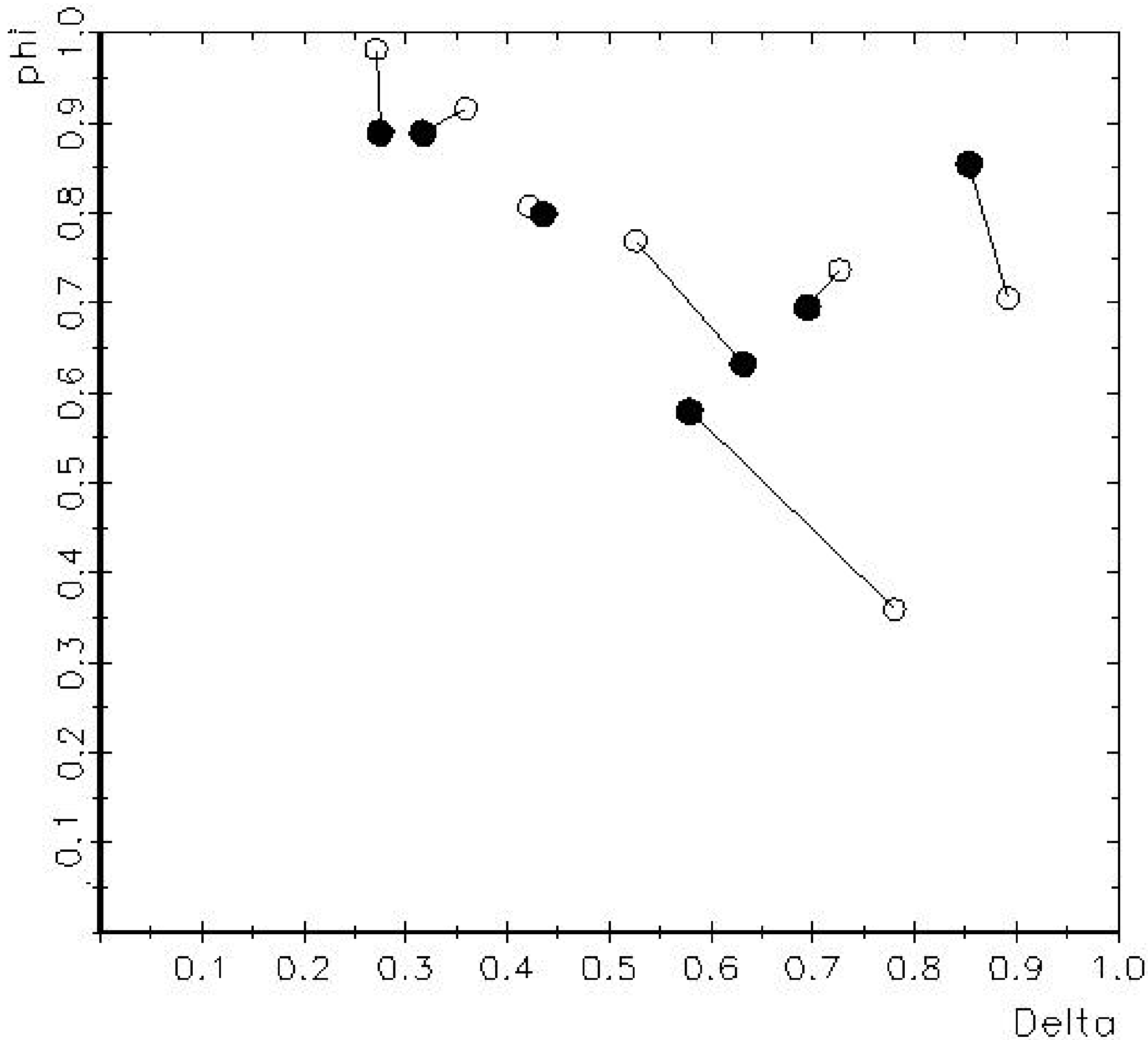

Figure 3 Transition Matrices for Patent Race

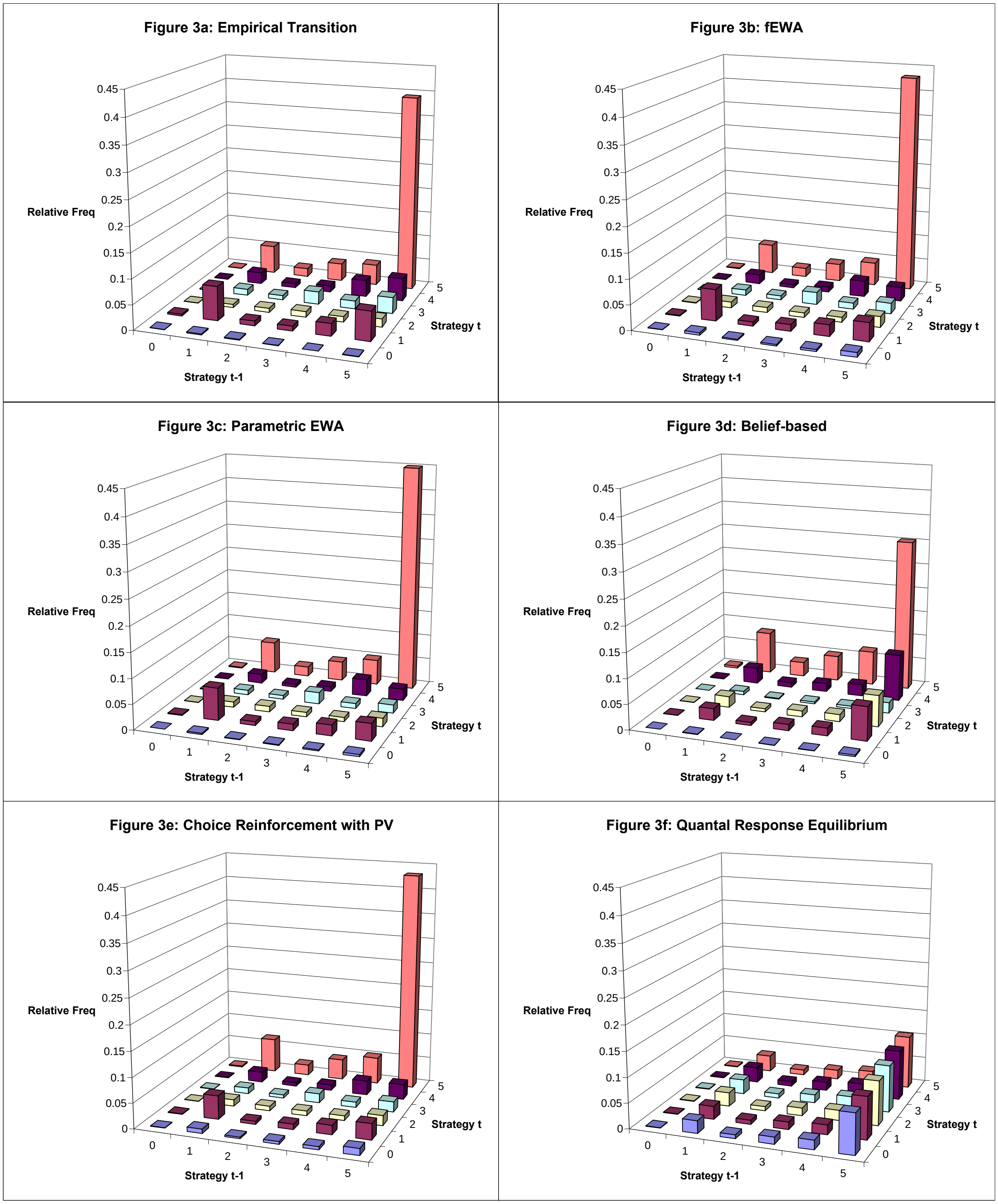

Figure 4: Empirical Frequency and Model Predictions for Continental Divide

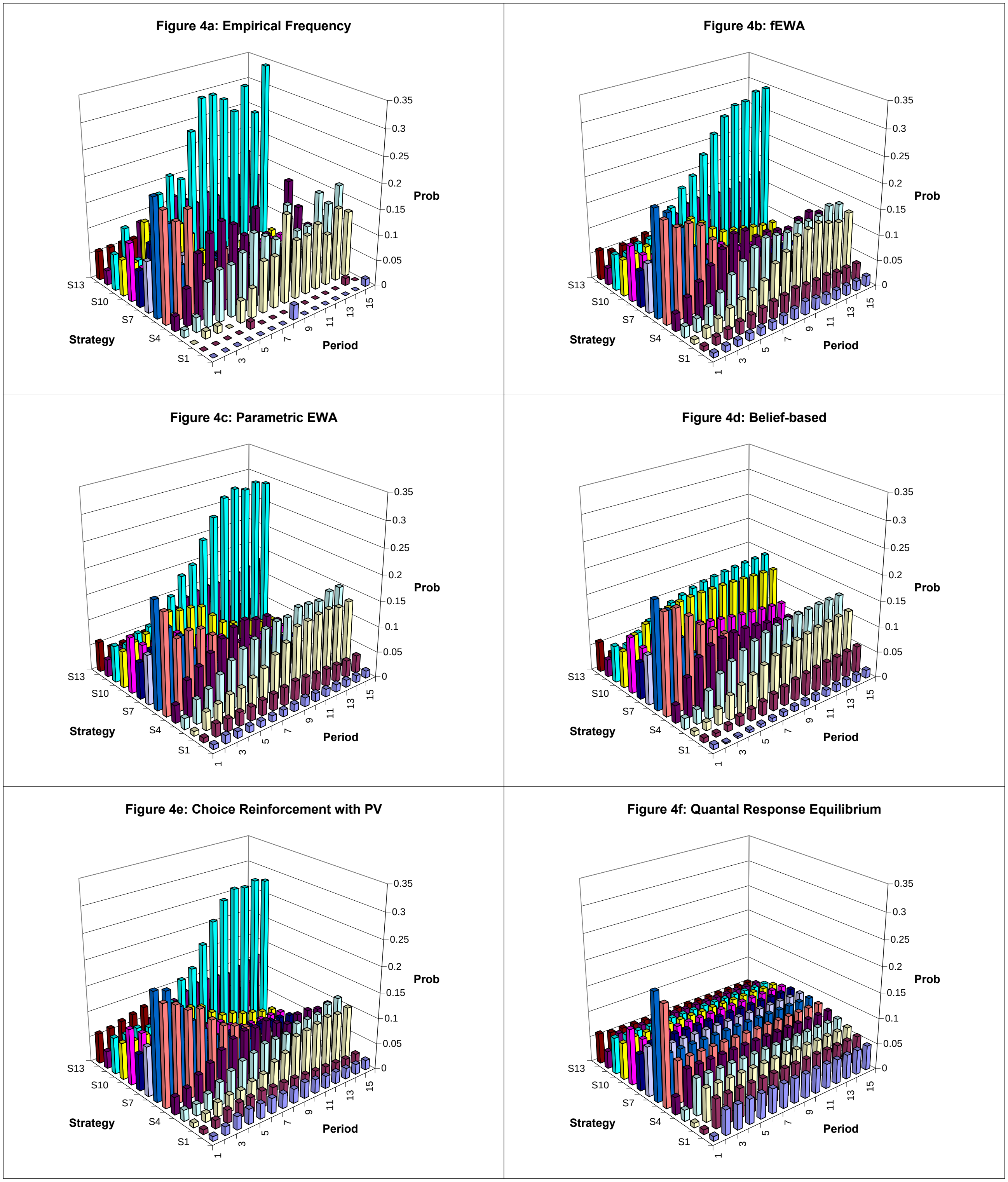

Figure 5: Empirical Frequency and Model Predictions for Traveler's Dilemma

Figure 5a: Empirical Frequency

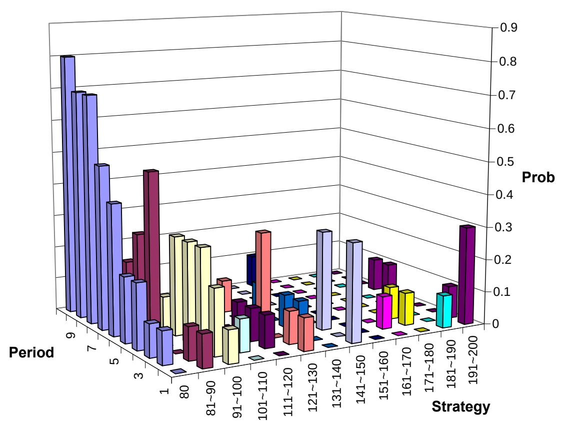

Figure 5b: fEWA

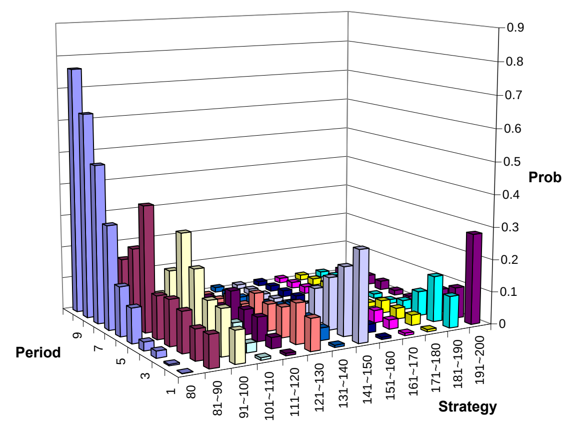

Figure 5c: Parametric EWA

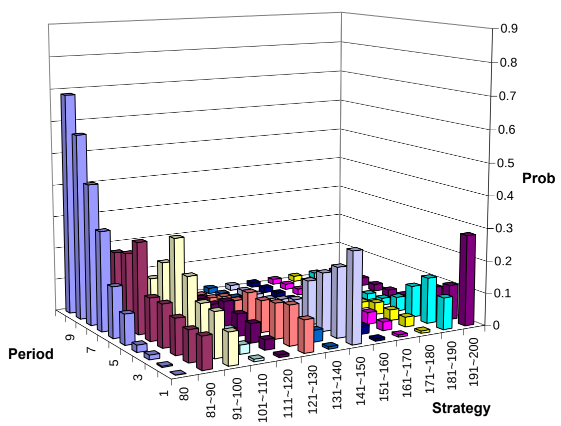

Figure 5d: Belief-based

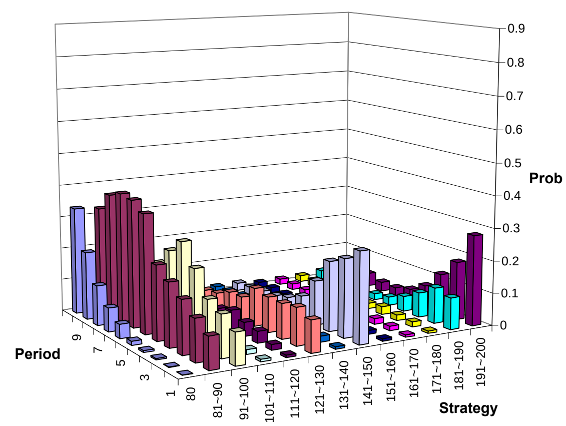

Figure 5e: Choice Reinforcement with PV

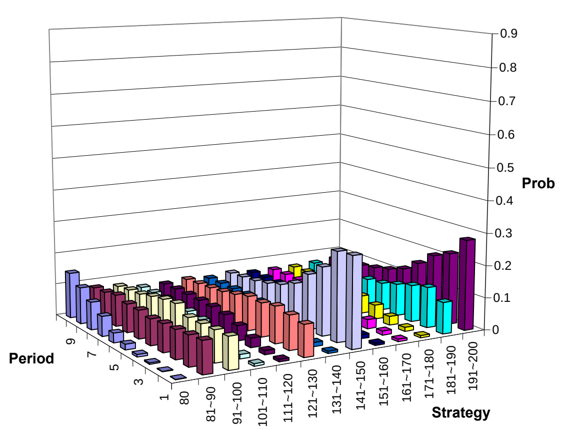

Figure 5f: Quantal Response Equilibrium