

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

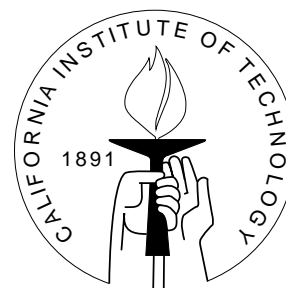
PASADENA, CALIFORNIA 91125

EWA LEARNING IN BILATERAL CALL MARKETS

Colin Camerer

David Hsia

Tech-Hua Ho



SOCIAL SCIENCE WORKING PAPER 1098

August 2000

EWA LEARNING IN BILATERAL CALL MARKETS

Colin F. Camerer
Division of Humanities and Social Sciences 228-77
Caltech
Pasadena CA 91125
camerer@hss.caltech.edu

David Hsia
University of Southern California

Teck-Hua Ho
Department of Marketing
Wharton School
University of Pennsylvania
Philadelphia PA 19104-6366
hoteck@marketing.wharton.upenn.edu

Abstract

This chapter extends the EWA learning model to bilateral call market games (also known as the "sealed-bid mechanism" in two-person bargaining). In these games, a buyer and seller independently draw private values from commonly-known distributions and submit bids. If the buyer's bid is above the seller's, they trade at the midpoint of the two bids; otherwise they don't trade. We apply EWA by assuming that players have value-dependent bidding strategies, and they partially generalize experience from one value/cost condition to another in response to the incentives from nonlinear optimal bid functions. The same learning model can be applied to other market institutions where subjects economize on learning by taking into consideration similarity between past experience and a new environment while still recognizing the difference in market incentives between them.

The chapter also presents a new application of EWA to a "continental divide" coordination game, and reviews 32 earlier studies comparing EWA, reinforcement, and belief learning. The application shows the advantages of a generalized adaptive model of behavior that includes elements of reinforcement, belief-based and direction learning as special cases at some cost of complexity for the benefit of generality and psychological appeal. It is a good foundation to build upon to extend our understanding of adaptive behavior in more general games and market institutions. In future work, we should investigate the similarity parameters, ψ and ω , to better characterize their magnitude and significance in different market institutions.

Keywords: Experimental economics, call markets, sealed-bid mechanism, learning

JEL Classification: D44, D83, C92

August 2, 2000. Thanks to Terry Daniel for supplying data. This research has been supported by NSF grant SBR 9730364 and a MacArthur Foundation Preferences Network postdoctoral fellowship to David Hsia.

1. Introduction

This paper is about learning in bilateral call markets. In these markets, buyers and sellers are privately informed of their values and submit their bids anonymously. If the buyer's bid is (weakly) more than the seller's ask, they trade at the midpoint of their bids. Understanding learning in bilateral call markets serves as a foundation for studying learning in more complex market institutions such as posted offers and double auctions. It also forces a generalization of learning models developed for simpler games to environments in which learning contingent on one realized random variable, such as a buyer's valuation in one trial, is generalized to similar valuations. Similarity-based generalization is a natural way to extend what is learned locally, which is undoubtedly important when people learn in very complex environments (in which learning has not thoroughly explored experimentally).

Studying learning in games is important because most of game theory revolves around the analysis of different equilibria of a wide range of games. The question of how, and whether, equilibrium actually arises was not thoroughly explored until recently. Now theorists actively study the dynamic properties of various models of evolution and learning, with a special focus on which types of equilibria these dynamics converge to (e.g., Weibull, 1998; Fudenberg and Levine, 1998). However, less attention has been paid to how well these dynamic models fit and predict data. The natural place to start is with data from laboratory experiments, where we have good control over the subjects' incentives and perceptions of the game they are playing, and the information they receive which helps them learn. Eventually, of course, these models should be extended to explain and predict learning in field settings (e.g., Ho and Chong, 1999).

Many of these models-- though not all-- have traditionally been classified into two groups: reinforcement, and belief learning. Reinforcement models update some unobserved propensity, reinforcement level, or attraction, according to what a chosen strategy actually earned (perhaps relative to some reference point, which might adapt). Belief models form beliefs based on some weighted average of previous observations of what other players have done. Beliefs are then used to compute expected payoffs of different strategies, and those with the highest expected payoffs are chosen with higher probability. Camerer and Ho (1999) created a more general model, experience-weighted attraction (EWA) learning, which hybridizes the main features of reinforcement and belief learning. In EWA, strategies have attractions which are updated by decaying lagged attractions and multiplying them by an experience weight (which is also updated), and adding either the payoff received from choosing a strategy, or δ times the foregone payoff from unchosen strategies. Attractions are then divided by a quasi-normalizing factor which controls whether attractions are averages or cumulations of past payoffs. More details of this model are given below.

This paper extends the EWA approach to bilateral call markets. In these markets, a single buyer and seller are each privately informed about their own value and cost, denoted V and C , respectively. The probability distributions of possible values and costs, $f(V)$ and $g(C)$, are commonly known. They submit bids, denoted $v(V)$ and $c(C)$. If the buyer is willing to pay at least as much as the seller demands, $v(V) \geq c(C)$, they trade at a price which is the midpoint of their bids, $(v(V)+c(C))/2$; otherwise they do not trade.

These are bilateral call markets because they are two-agent examples of general call markets. In call markets, many traders submit demand and supply schedules, and the market is "called" at the price where supply meets demand. The mechanism used to determine prices in the two-agent case is also called the sealed-bid mechanism.

We investigate bilateral call markets for two reasons. First, there is an enormous amount of experimental evidence about the way in which market institutions influence prices and quantities, and the speed and nature of convergence, when there are many traders (see Holt, 1995, for a review). Very few of the learning models developed to carefully explore learning in simpler strategic games have been applied to market data. The only exceptions we know of are two studies of call markets with many buyers and sellers. Cason and Friedman (1999) applied a form of direction learning to a laboratory call market with four buyers and four sellers and random supply and demand each period. They found that partial adjustment toward ex post optimal bids explained a significant portion of the observed behavior. In addition, there is strong asymmetry in the adjustment and significant 'observational learning' taking place. The result shows that actual received payoffs are not the only information traders use to guide their decisions. Hsia (1999) applied EWA (and some models which are extreme special cases of it) to the same data and found that EWA did significantly better than both reinforcement and belief based models.

In the call markets studied by Cason, Friedman, and Hsia, Bayesian Nash equilibrium strategies require markdowns or markups of valuations (i.e., a buyer should bid some fraction of her value) which are approximately constant for all valuations, and small compared to the bilateral call markets. This relatively weak incentive to optimize may have contributed to the asymmetry in traders' adjustment and the inability to get significant learning parameters in some cases. By extending EWA from games to bilateral call markets, we hope to learn something about how the model can be applied to more complicated market institutions, in which bidding strategies are complex (nonlinear) and there is a torrent of information (bids, asks, and acceptances) which influences learning.

Second, it is cognitively plausible to think that when a buyer draws a valuation V , bids $v(V)$, and either makes a trade or doesn't, she learns something about how she should bid for valuations similar to V . This effect can be captured in a model which "learns less" about events that are far from V and $v(V)$, or less similar to V and $v(V)$, in a sense that we make precise later. Similarity-based generalization of learning is well-established in cognitive psychology and heuristics like these are used in machine learning in computer science (see Kaelbling, Littman, and Moore, 1996, for a summary). Similarity-based learning is also arguably a "cognitively economical" heuristic because scarce attention is allocated where it is likely to be most useful-- namely, in the vicinity of the current valuation and bid. Our approach revives the generalization parameter in Roth and Erev (1995) and is also closely related to the approach of Sarin and Vahid (1999). The latter showed that "spillover" of reinforcement from a chosen strategy to neighboring strategies could explain why players actually converge to equilibrium much more rapidly than they are predicted to solely by choice reinforcement, in order-statistic experiments conducted by Van Huyck, Battalio and Rankin (1996).

The long-run goal of the approach we take in this paper is a general theory of learning in markets and other complex economic environments. Such a theory would be a

scientific breakthrough if it could connect game-theoretic learning models to approaches in psychology and computer science. A theory of market learning would also be useful in practice. Theorists are now actively engaged in designing more and more complex incentive systems, or "mechanisms", to help societies and firms incentivize agents to act in ways which maximize some objective function. Business firms are also designing more ambitious trading systems-- e.g., internet auction companies Ebay and Ubid. These mechanisms are usually discovered by trial-and-error learning or managerial hunches, or by theorists searching for sets of rules which balance "incentive compatibility" and "individual rationality". However, the definition of rationality in mechanism design ignores the cognitive difficulty of computing how to behave under the rules, or the difficulty of learning from experience. In practice, every applied mechanism designer has an intuitive sense that some mechanisms work beautifully in theory, but will fail in practice because they are too opaque for agents to figure out or learn (e.g., Ledyard, 1993). A theory of learning could formalize the designers' intuitions by suggesting precise "learnability" constraints. Adding these constraints would inform which mechanisms would be easiest to learn, and hence most likely to work well in practice (e.g., Friedman, 1998).

The paper proceeds as follows. The next section (2) describes the EWA approach, briefly summarizes earlier results (and alternative theories), and illustrates the application of EWA to a "continental divide" coordination game. Section 3 describes previous results on bilateral call markets (also called the "sealed-bid mechanism").

2. Previous research on learning in games

There have been dozens of studies fitting different learning models to experimental data. Rather than summarize this large and growing literature, we will only mention some highlights and dwell on precursors which motivate our analysis of the call market data. We start with notation. We study n -person normal-form games. Players are indexed by i ($i=1, \dots, n$), and the strategy space of player i , S_i consists of m_i discrete choices, that is, $S_i = \{s_i^1, s_i^2, \dots, s_i^{m_i-1}, s_i^{m_i}\}$. $S = S_1 \times \dots \times S_n$ is the Cartesian product of the individual strategy spaces and is the strategy space of the game. $s_i \in S_i$ denotes a strategy of player i , and is therefore an element of S_i . $s = (s_1, \dots, s_n) \in S$ is a strategy combination, and it consists of n strategies, one for each player. $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ is a strategy combination of all players except i . S_{-i} has a cardinality of $m_{-i} = \prod_{j=1, j \neq i}^n m_j$. The scalar-valued payoff function of player i is $\pi_i(s_i, s_{-i})$. Denote the actual strategy chosen by player i in period t by $s_i(t)$, and the strategy (vector) chosen by all other players by $s_{-i}(t)$. Denote player i 's payoff in a period t by $\pi_i(s_i(t), s_{-i}(t))$. Many theories assume that each strategy has a numerical attraction, which determines the probability of choosing that strategy. Learning consists of changes in attractions based on experience. (In "rule learning" the strategies are rules which map history into choices, and players update the attractions of these rules; e.g., Salmon, 199x and Stahl, in press).

Learning models of this sort require a specification of initial attractions, how attractions are updated by experience, and how choice probabilities depend on attractions. The core of the EWA model is two variables which are updated after each round. The first variable is $N(t)$,

which we interpret as the number of “observation-equivalents” of past experience. The second variable is $A_i^j(t)$, player i 's attraction of strategy j after period t has taken place.

Updating is governed by two rules. The experience weight $N(t)$ is updated according to $N(t) = \phi(1 - \kappa)N(t-1) + 1$. Attractions are updated according to

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

Where $I(x, y)$ is an indicator function which is one when $x = y$ and 0 otherwise. Intuitively, attractions are equal to decayed (by ϕ), experience-weighted ($N(t-1)$) lagged attractions, plus reinforcement for the received payoff (if $s_i^j = s_{i(t)}$) or δ times the reinforcement for a foregone payoff.¹ This numerator is then quasi-normalized by dividing by $\phi(1 - \kappa)N(t-1) + 1$. The free parameters δ , ϕ , and κ all have intuitive interpretations.

The weight on foregone payoffs δ is a kind of “imagination” of foregone payoffs, or “simulation” of outcomes under alternative competitive scenarios; it might also be considered responsiveness to opportunity costs or regret. The decay parameter ϕ is naturally interpreted as the degree to which players realize other players are adapting, so that old observations on what others did become less and less useful. (In current work we are trying to endogenize ϕ so that it need not be estimated from the data, and can change throughout an experimental session as players sense that the strategic environment is becoming more or less stable.) The parameter κ determines the growth rate of attractions, which reflects how quickly players lock in to a strategy or, in machine learning terms, how quickly players shift from “exploring” an environment to “exploiting” what they have learned. When $\kappa = 0$, attractions are weighted averages of lagged attractions and past (δ -weighted) payoffs, where the averaging weights are $\phi N(t-1) / (\phi N(t-1) + 1)$ and $1 / (\phi N(t-1) + 1)$. When $\kappa = 1$, the denominator becomes one and attractions are (decayed) cumulations of past payoffs. When attractions are cumulations, a strategy which is chosen frequently and yields positive payoffs can build up a large “lead” over unchosen strategies, so that exploration is brief and a player quickly turns to exploiting their historical information by locking in to one strategy rapidly.

Note that while we have not subscripted the key parameters δ , κ , and ϕ , they obviously could be different across players or games. Attractions must determine probabilities of

choosing strategies in some way. We generally use the logit form, $P_i^j(a, t+1) = \frac{e^{\lambda \cdot A_i^j(a, t)}}{\sum_{k=1}^m e^{\lambda \cdot A_i^k(a, t)}}$,

but a power form fits about equally well (Camerer and Ho, 1999a).

Figure 1 shows a cube with axes representing the imagination parameter δ , the change parameter ϕ , and exploration/exploitation parameter κ . Many existing theories are simply extreme cases of EWA learning which are represented by points or edges of the cube. For example, cumulative reinforcement, average reinforcement, weighted fictitious play are edges and Cournot and fictitious play are vertices of this cube.

When $\delta=0, \rho=0$ (and $N(0)=1$), then $N(t)=1$ and the attraction updating equation becomes $A_i^j(a, t) = \phi \bullet N(t-1) \bullet A_i^j(a, t-1) + [\delta + (1-\delta)] \bullet I(s_i^j, s_i(t)) \bullet \pi_i(s_i^j(t), s_{-i}(t))$. This is the simplest form of cumulative choice reinforcement (e.g., Roth and Erev, 1995, with some features left out). When $\delta=0, \rho=\phi$ (and $N(0)=1/(1-\rho)$), the attraction updating equation becomes $A_i^j(a, t) = \phi \bullet N(t-1) \bullet A_i^j(a, t-1) + (1-\phi)[\delta + (1-\delta)] \bullet I(s_i^j, s_i(t)) \bullet \pi_i(s_i^j(t), s_{-i}(t))$. This is a form of averaged choice reinforcement (attractions are averages of previous attractions and incremental reinforcement). The most surprising extreme case is weighted fictitious play. When $\delta=1, \phi=\rho$, then the attractions are updated according to

$$A_i^j(a, t) = \frac{\phi \bullet N(t-1) \bullet A_i^j(a, t-1) + \pi_i(s_i^j(t), s_{-i}(t))}{\phi \bullet N(t-1) + 1}$$

This updating rule corresponds exactly to updating of expected payoffs according to “weighted fictitious play”, in which players' weight the last observation of what others did by one, and from t periods ago by $\phi^{(t-1)}$, form a belief which is a normalized average of these weighted observations, and use those beliefs to calculate expected payoffs. Weighted fictitious play includes fictitious play ($\phi=1$) and Cournot best-response dynamics ($\phi=0$) as special cases.

Seen as vertices of the EWA parameter cube, it is obvious that reinforcement and belief learning models are closely related rather than fundamentally different (as many people have suggested). Belief learning (of the weighted fictitious play sort) is simply a kind of generalized reinforcement in which unchosen strategies are reinforced by foregone payoffs, as strongly as chosen strategies are, and reinforcements are averages rather than cumulations.

The EWA cube also suggests that there is no good *empirical* reason to think that players' parameter configurations would necessarily cluster on the vertices corresponding to reinforcement and belief learning. Indeed, early studies did not consider a wide range of parameter values and hence, have never found clustering on particular vertices. (Instead, the typical study simply chose an edge or vertex and asked how well that specific learning rule matched the data, compared to a benchmark like Nash equilibrium or random choice.) Furthermore, the intuition behind reinforcement is compelling-- namely, chosen strategies are reinforced more strongly-- and the intuition behind belief learning is compelling too-- unchosen strategies should be strongly reinforced too. There is no scientific reason to choose one intuition or the other, when one can respect both intuitions with a value of δ between 0 and 1 (i.e., a learning model in the interior of the cube).

Continental divide game

To illustrate how EWA and the other learning models fit, it is useful to look briefly at a specific game-- the “continental divide game” (Van Huyck, Cook, and Battalio, 199x). In this game, if the median starts at 7 or less and subjects best-respond, they will eventually work their way to a pure-strategy equilibrium at 3. If the median starts at 8 or above, however, best-responding will eventually converge to an equilibrium of 12. Both equilibrium payoffs are shown in italics. The payoff at 3 is about half as much as at 12.

Table 1: Payoffs in ‘continental divide’ experiment, VHCB (in press)

choice	Median 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

This game captures the possibility of extreme sensitivity to initial conditions (or path-dependence), which has recently begun to fascinate all sorts of scientists interested in chaotic dynamics and complex systems. Path-dependence is often evident in physical systems. I once went mountain climbing in Alaska with a friend. We stood on the continental divide, which is the imaginary line (marked on a map) which marks the point at which the direction of water flow changes from one direction to the opposite. We poured water from a canteen right at the divide; some water trickled south and some trickled north. Eventually, the north-flowing water made it to the Arctic Ocean, and the south-flowing water to the Pacific. Molecules that began imperceptibly close together ended up a thousand miles apart.

What happens in the continental divide game? Figure 2 shows time series of median choices from ten sessions conducted by VHCB. Each line represents the time series of medians from a separate group. An imaginary line between choices 7 and 8 acts precisely like the continental divide. Groups that started with medians of 8 or above are inexorably drawn to 12-13; groups that started with low medians converge to 3-6. Remember that the low groups, in equilibrium, earn about half as much as the high groups do. Tiny historical accidents have large, persistent earnings consequences.

We estimated the EWA model on the continental divide data by fixing the initial attractions so that predicted choice frequencies in period 1 match the data exactly. We then used maximum-likelihood estimation to find parameters which fit the first 10 periods best, then used those parameters to forecast behavior in the last five periods. The procedure uses the actual history each subject faced (i.e., their own choices, the medians in their group) so that each person's predicted choice probabilities are different in each period. Table 2 shows the estimated parameters for EWA learning model and for restricted cases that correspond to reinforcement ($\delta=0, \rho=0, N(0)=1$) and belief based ($\delta=1, \rho=\phi$) models.

Table 2: The estimated parameters (standard errors) for various learning models.

	EWA Learning	Reinforcement	Belief Learning
<i>a1</i>	<u>0.0000</u>	<u>0.0000</u>	<u>0.0000</u>
<i>a2</i>	0.0000	0.0000	0.0000
<i>a3</i>	0.0000	0.0000	0.0000
<i>a4</i>	0.0000	0.0000	0.0000
<i>a5</i>	0.6152	0.6240	0.4965
<i>a6</i>	0.9304	1.2916	0.8478
<i>a7</i>	0.9292	1.2053	0.8602
<i>a8</i>	0.7781	0.8186	0.7319
<i>a9</i>	0.7052	0.7064	0.6907
<i>a10</i>	0.7230	0.7628	0.7618
<i>a11</i>	0.6604	0.6677	0.6839
<i>a12</i>	0.7432	0.7519	0.6963
<i>a13</i>	0.5016	0.4326	0.5404
<i>a14</i>	0.7017	0.5873	0.6575
ϕ	0.6112	0.5630	<u>0.0008</u>
δ	0.7495	<u>0.0000</u>	<u>1.0000</u>
ρ	0.0008	<u>0.0000</u>	<u>0.0008</u>
$N(0)$	0.7621	<u>1.0000</u>	0.7722
λ	4.5210	2.0852	6.5572
-LL	-1189.85	-1438.66	-1369.94

Figure 3 shows the actual frequencies of strategies 1-14, pooled across all 10 cohorts for all 15 periods. This plot masks the dramatic path-dependence evident in Figure 2, but shows the general tendency toward bifurcation (and also shows that convergence toward the high-number equilibrium at 12 is sharper than toward the low-number, Pareto-dominated equilibrium).

As a comparison, we simulated choices using the estimated EWA parameters. Starting with initial attractions, we generated first period choices stochastically using the probability profile from the attractions that are identical across individuals. As subjects update their attractions using actual and hypothetical payoffs, attractions of each individual become different depending on the actual choices. Thus in the subsequent periods, subject will choose strategies with different probabilities. We simulated the continental divide games using all three learning models with the estimated parameters in Table 2. Figure 4 compares the distribution of choices across strategies for 1000 simulations pooled across all players and periods. We can see that none of the models can quite match the sharpness of the actual frequencies, with the reinforcement learning model doing the worse, completely missing the critical features of the game. Belief learning model did much better but adjusted too quickly on the lower equilibrium, resulting in a lower peak in the distribution at 3 compared to 5 from the data. EWA learning did a little better still, though missing on the sluggish side with the lower peak around 5 and 6.

Figure 5 shows the paths of 10 typical medians from the simulation using EWA learning model. These results show why it is potentially important to hybridize features of learning models separately. The best-fitting model combines the responsiveness of belief learning

(high δ) with the cumulation of reinforcement models (high κ), because players quickly lock in, exploiting their learning from the first few trials. Simply averaging expected payoffs and reinforcement levels would not produce this kind of gene-splice or hybridization.

Figure 6 shows similar paths from the belief learning restriction ($\delta=1, \kappa=0$) and Figure 7 shows paths from the choice reinforcement restriction ($\delta=0, \kappa=1$). Reinforcement fits very poorly. (Reinforcement will generally do poorly when the set of strategies which are chosen at the beginning of a game are different than the strategies which are chosen toward the end, as in "market games" (proposer competition), see Roth and Erev, 1995, or beauty contest games see Camerer and Ho, 1999.) Belief learning fits rather well. EWA does slightly better than belief learning in two small ways, however: (i) Notice that strategies 1-4 are never chosen in early periods, but are frequently chosen in later periods; and (ii) 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 second-order 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). EWA does so rather well; belief learning does not. Thus, in this particular sample, while EWA does not add much visually to belief learning, the crucial parameter estimates are significantly different than the belief learning restriction, and belief learning is not able to accelerate and brake as quickly as EWA.

The continental divide data illustrate how EWA can improve marginally on belief learning, and substantially on reinforcement learning, by hybridizing their features into a novel mixture. More generally, empirical results comparing reinforcement and belief models have been mixed. Many direct comparisons of the two favored reinforcement models (Mookerjee and Sopher, 1994, 1997; Erev and Roth, 1998). However, when weighted linear combinations of reinforcements and expected payoffs are used to fit data, the weight on expected payoffs is about ten times higher than on reinforcements (Erev and Roth, 1998, p. 867; Battalio, Samuelson and Van Huyck, 1999; Munro, 1999). The mixed results might be due to use of games and techniques which do not have enough power to distinguish between models (Salmon, 1999).

Studies which estimate the EWA model provide an easy way to compare the extreme special cases to one another, and to EWA, by testing the implied restrictions on parameter values. Table 3a summarizes results of model estimation from 32 data sets, including our earlier papers, call markets (Hsia, 1998), cost allocation processes (Chen, 1998), extensive-form centipede games (Camerer, Ho and Wang, 1999), "unprofitable" games (Morgan and Sefton, 1998), signaling games (Anderson and Camerer, 1999), patent race games with iteratively dominated strategies (Almadoss and Rapoport, 1999), pricing games (Abramson, 1998), and 5x5 matrix games (Stahl, 1999).

The goodness-of-fit statistic, summarized in Table 3b, is -1 times log likelihood except in Chen (1999). The column "EWA" reports the $-LL$ of the EWA model. The reinforcement and belief models report the *difference* between the $-LL$'s of those models and the EWA statistic. (Positive differences mean that EWA fits better.)

Values of δ tend to be between .5 and 1 in most studies except those in which games have only mixed-strategy equilibria, where δ is close to zero. The value of ϕ is reliably around .9 or so, with a couple of exceptions.

What about model comparisons? The fairest comparisons estimate parameters on part of a sample of data and forecast choices out-of-sample, so that models with more parameters will not necessarily fit better.² (Indeed, if they succeed in-sample by overfitting, they will do particularly badly when predicting out-of-sample.) In those 11 comparisons (denoted “OUT” in the third column from the right), EWA predicts better than reinforcement in every case, although usually modestly. EWA predicts better than belief learning in 9 of 11 cases, by a large margin in some data sets, and does worse in two constant-sum games.

Of course, EWA necessarily fits better in the other 20 in-sample comparisons than reinforcement and belief models because the latter are special cases. But it also does better in almost all cases when penalizing EWA for extra degrees of freedom using a standard ψ^2 test, or the Akaike or Bayesian criteria. For example, if the difference in LL is 4 points or more then the special-case restriction will be rejected by the ψ^2 test. By this criterion, EWA fits more accurately than belief learning in all in-sample comparisons, and fits more accurately than reinforcement in 16 out of 20 comparisons.

Figure 1 shows estimates of EWA parameters for 20 of the 32 studies in Table 3a.³ Each point represents the results of one study. There is no strong tendency for parameter estimates to cluster in any particular corner or vertex, although there are quite a few points in the lower corner corresponding to averaged-reinforcement learning ($\delta=0$, $\kappa=0$) with high ϕ . We are not sure why points are different for different games, although we note that every study which has looked for cross-game variation has found statistically significant variation. Eventually, of course, it would be good to have a theory which specifies parameter values in advance, from empirical regularity or the structural properties of a game, and we are currently researching this question.

Given the results from almost two dozen studies shown in Figure 1, and the usefulness of a model which combines the intuitions of reinforcement and belief learning, it is hard to think of a good empirical reason why attention should continue to be focussed on any extreme special case theory rather than on EWA.

3. Bilateral call market experiments

There have been several experiments on bilateral call markets (see Roth, 1995, p. 253-348; Camerer, in progress, for reviews.) The earliest work is by Radner and Schotter (1989). They studied the case in which values and costs are uniformly distributed from 0 to 100. In this case, there are equilibrium bid functions for risk-neutral agents which are piecewise linear, and require buyers to underbid their values and sellers to overbid their costs. For example, the buyer should bid her value ($v(V)=V$) up to 25, and bid $v(V)=(25+2V)/3$ for higher values. In some sessions they use nonuniform value distributions which predict bid functions with lower slopes. They find that empirical bid functions, estimated by regressions of actual bids on values, have slopes and slope

differences (reflecting the piecewise linearity) which are rather close to those predicted by theory, and which change with the value distribution in the predicted way. Radner, Schotter and Zheng (check cite) extended their study by forcing players to submit reservation prices to agents, who bargained on their behalf.

Radner and Schotter found that when players bargained face-to-face, rather than by submitting bids, that efficiency was quite high. In fact, efficiency was higher than predicted by any equilibrium bid functions, so the players were violating individual rationality constraints for the sake of joint gain. The variance of prices was also high (compared to the sealed-bid control), which suggests that some players were much better bargainers than others and were able to put their skill to use when bargaining face-to-face. Players were able to jointly create more surplus than predicted by theory. Valley et al (1998) unpacked this finding by comparing the sealed-bid mechanism with treatments in which players passed written bids back and forth, and communicated face-to-face, before submitting sealed bids. They found that while subject pairs did not typically reveal both of their values to one during the pre-bidding communication, they were frequently able to coordinate on a single price, which they both bid. As in Radner and Schotter, efficiency was higher than predicted by any equilibrium in the communication conditions, particularly for face-to-face communication.

Several studies by Rapoport and colleagues extended the work of Radner and Schotter (using no pre-bid communication). Rapoport and Fuller (1995 used the “strategy method”, eliciting the entire $v(V)$ function. They used symmetric uniform $[0,100]$ values in one experiment. In a second experiment the seller's cost was uniform from $[0,100]$ but the buyer's value was uniform from $[0,200]$. In that case, there is a piecewise linear equilibrium in which buyers make a constant bid of 116.7 for values above 150. From a learning point of view, these bid functions with flat portions are interesting because it is an open question whether bidders can learn to bid a constant rather than a function which increases monotonically with value.

Rapoport and Fuller found reasonable conformity with the piecewise-linear functions, except that subjects tended toward more full revelation of values over time (i.e., in the opposite direction of equilibrium).

Daniel, Seale, and Rapoport replicated the asymmetric-value experiment with similar results.¹ They also ran two experiments in which the seller's costs are uniforms from $[0,100]$ and $[0,20]$ respectively, and the buyer's value is uniform from $[0,200]$. We analyze their data in experiment 2 because the equilibrium bid functions are farthest from fully-revealing bidding, so there is more room to observe learning. The extreme asymmetry in values produces a (piecewise) linear equilibrium solution (LES) of

¹ Rapoport, Daniel and Seale (1999) replicated the DSR experiment 1. They also ran an experiment in which the buyer's values are uniform $[100,200]$ and seller costs are uniform from $[0,200]$. These parameters flip around the bid functions so they are the opposite of those (i.e., buyer and seller bid functions switch) from experiment 1.

$$B = \begin{cases} V & \text{if } B < 50 \\ 50/3 + (2/3)V & \text{if } 50 < B \leq 70 \\ 190/3 & \text{if } B > 70 \end{cases} \quad \text{for all } V. \text{ (buyers)}$$

and

$$A = 50 + (2/3)C \quad \text{for all } C. \text{ (Sellers)}$$

where V and B are the buyers' values and bids and C and A are the sellers' costs and asks.

Most call market experiments have been done with symmetric uniform distributions, which produce linear equilibrium bid functions as Bayesian Nash equilibrium bidding strategies. (There usually are other nonlinear equilibria, such as "one-price" equilibria and others.) When we usually observe approximately linear bid functions in these markets, it is not clear whether traders "learned" to bid linearly or just happen to use a constant markdown ratio as a convenient strategy. The extreme asymmetry of the linear equilibrium in this case provides the buyers with the opportunity and incentive to learn a more complicated strategic behavior, since the optimal markdown ratios vary from 0% to almost 70% over the range of possible values.

In their experiment, there were 10 pairs of randomly matched buyers and sellers each period. They were rematched each period but did not change roles as buyers or sellers. The same sets of 50 values and costs were used in different order for each of the 10 pairs. Figure 8 shows all the bids in experiment 2 compared to the linear equilibrium and Figure 9 shows the same bids over time as bidders modify their behavior. There is clearly a change in bidding behavior over time as buyers learn the optimal bid function. Next we will look at buyers' learning behavior in this market.

4. Similarity-based generalization of EWA learning

We are interested in using EWA learning as a more general alternative to adaptive behavior in this market. In contrast to DSR's direction learning model to be discussed later, EWA learning does not a priori assume a direction of adjustment for each outcome, nor an identical adjustment for all values. Rather it considers the foregone payoffs for all possible value/strategy combinations and updates the propensities for each strategy individually depending on the local impact of a choice. As we will see later, EWA model does a comparable job in fitting the data at a cost of complexity for the sake of generality.

To use EWA learning model, we first discretize the values into 10 equal increments from 0 to 200. Corresponding to the middle of each value increment, there are 16 strategies of evenly spaced markdown ratios from 0% to 75%. Thus we have 10 values with 16 possible strategies each, making a total of 160 value dependent strategies. Assuming that the value dependent initial attractions for a bidder i are given by normal distributions of given means and variances, they can be written as, $A_i^{b/v}(0) = v \max^* v * N(b_i(0), \mu_i | v, \sigma_i | v)$, where v is the random value draw, $\mu_i/v = \alpha_i + \beta_i v + \gamma_i v^2$, the mean absolute markdown conditional on realized values, and $\sigma_i/v = \eta_i v$, the variance of absolute markdowns conditional on realized values, or

$$A_i^{b/v}(0) = v \max^* v * \frac{1}{\sqrt{2\pi\eta_i v}} \exp\left\{-\frac{[b_i(0) - (\alpha_i + \beta_i v + \gamma_i v^2)]^2}{2(\eta_i v)^2}\right\} \quad (1)$$

In this particular specification, the distribution of average markdown (in absolute price rather than in percentage) is assumed to be a quadratic (which includes linear bid function as a special case) function of values. The initial attractions are assumed to be normally distributed around this mean at each of the 10 value with the standard deviation proportional to the private values for that distribution. Thus we scale the normal probability density function by private value, v , to put initial attraction for different values on similar absolute levels. The parameter v_{\max} scales initial attractions to maximize the likelihood function in a way similar to setting an initial attraction to zero to achieve identification in earlier models.

Given the initial attractions, attractions are updated according to,

$$A_i^{b/v}(t) = \frac{\phi^\tau N(t-1) \cdot A_i^{b/v}(t-1) + \tau \pi_i^{b/v}(b_i, v_i, b_{-i}(t))}{\phi^\tau (1-\kappa)^\tau \cdot N(t-1) + \tau}, \quad t \geq 1 \quad (2)$$

$$N_i^{b/v}(t) = \phi^\tau (1-\kappa)^\tau \cdot N_i^{b/v}(t-1) + \tau, \quad t \geq 1 \quad (3)$$

for subject $i = 1$ to n , strategy $S^{b/v}$ and value draw v .

where $\tau = e^{-\psi|v-v_i(t)|-\omega|b-b_i(t)|}$

Parameters ϕ , $N_i^{b/v}(t)$ and κ have the same interpretations as in Camerer and Ho, denoting discount in attractions, experience equivalent count and whether the experience weight depreciates faster than the attractions respectively. In Eqn. 3, the experience weight for each value dependent strategy is increased by τ each period, rather than by 1 in Camerer and Ho to reflect the partial generalization of experience to neighboring values and strategies. The parameter τ takes on values between 0 and 1 and indicates the “similarity” between strategies. The variables ψ and ω decay the spillover of experience from values and bids using the distances between realized and hypothetical values and bids, $v-v_i(t)$ and $b-b_i(t)$, respectively. The distance between bids measures the “closeness” in strategies and the distance between values measures the “similarity” in experience. In bilateral call markets, they both take the form of difference in prices. The two parameters allow us to separate the effect of the two distances qualitatively. For example, τ for a chosen value/strategy combination will be 1 since both $v-v_i(t)$ and $b-b_i(t)$ are equal to zero. On the other hand, if there is little generalization in values or strategies for a particular game, both ψ and ω will be relatively large, causing τ to fall very quickly away from the chosen value/strategy pair, resulting in little increase in the experience count.

Foregone payoffs, $\pi_i^{b/v}$, are easily calculated by profits from hypothetical bids since a buyer knows the opposing seller’s ask from the price after the market closes. Finally we use

the logit model to convert the attractions for each strategy to probabilities of choosing one, given by, $P_i^{b/v}(t) = \frac{e^{\lambda_i^{b/v}(t)}}{\sum_{k=1}^{m_i} e^{\lambda_k^b(t)}}$.

Seller's Adaptive Behavior

The sellers in DSR's experiment also appear to modify their asks over time as shown in Figure 10, although the change is not as dramatic as in buyers' case. We use the same EWA learning developed for the buyers to study sellers' adaptive behavior to further generalize the application of EWA learning model to different learning environments. The sellers' cost are drawn from a uniform (0,20) and observed asks are in the range of (9, 100) with one ask at 150. The discrete costs in the models are (2, 6, 10, 14, 18) and the strategies are (5, 10, 15, 20,..., 85, 90, 95, 100). This evenly spaced 5 by 20 grid covers all the value/strategy combinations in the data except one at 150, which is approximated in the estimation by the closest strategy at 100. The sellers' strategies are simply prices, rather than absolute markups as in the case of buyers for simplicity. An absolute markup over cost should produce very similar results since the range of costs is from 0 to 20 and the optimal ask function is $A = 50 + 2/3 * (\text{cost})$, resulting in a maximum difference of 13.3 in strategy space based on the optimal bid function. On the other hand, a model of markup ratios will miss many asks unless a very large number of strategies are used in the model because the intercept of the optimal ask function and the limited range. In addition, players did not appear to use markup ratio as strategies in the experiment.

The initial attraction profile is modeled as normal distributions centered around a linear function defined by $(\alpha + \beta * \text{cost})$. A linear function rather than a quadratic is used for simplicity since the optimal bid function is linear and there is little evidence of quadratic bid function by the sellers in the experiment. V_{\max} scales the initial attraction from standard normal distributions with standard deviations estimated by the optimization procedure as noted before.

Given the initial attractions, attractions for a strategy, $S^{a/c}$, are updated just like the

buyers before according to, $A_i^{a/c}(t) = \frac{\phi^\tau N(t-1) \cdot A_i^{a/c}(t-1) + \tau \pi_i^{a/c}(a_i, c_i, a_{-i}(t))}{\phi^\tau (1-\kappa)^\tau \cdot N(t-1) + \tau}$, $t \geq 1$

and, $N_i^{a/c}(t) = \phi^\tau (1-\kappa)^\tau \cdot N_i^{a/c}(t-1) + \tau$, $t \geq 1$

where subject $i = 1$ to n , with strategy $S^{a/c}$ for an ask, a , given a cost draw c , and

$$\tau = e^{-\psi|c-c_i(t)| - \omega|a-a_i(t)|}$$

The parameter ψ is the decay for cost distance and ω is the decay for ask distance. Together they measure the spillover or generalization of learning from one ask/cost combination to another similar to the buyers' model. All the other parameters have similar interpretations as corresponding ones in the buyers' model.

Estimation Results

The nonlinear optimization was done with SAS Proc NLP using a large number of starting points to ensure a global maximum was achieved. Table 4 shows estimated parameters for experiment 2 using a one-segment and a two-segment EWA learning model. We focus on the results for experiment2 because of the greater asymmetry in value distributions produce stronger incentive for a nonlinear bid function and more dramatic adaptive behaviors. Estimated parameters for ψ and ω are quite small, indicating significant spillover of reinforcement to other values and strategies. A typical distribution of the generalization parameter, τ , is shown in Figure 11 for a value draw of 100 and a markdown ratio of 30%. The value of τ falls off from $\tau=1$ quickly as we move away from the realized value/choice combination but there is still substantial learning for some strategies with values as low as 50. There is no discounting of attractions ($\phi=1$) and a small κ means that choices are based more on average attractions rather than cumulative ones. The two-segment model seems to fit the data better but does not produce noticeably different initial attractions between the two segments. Table 5 gives the bootstrapped standard errors of estimated parameters. They are generally smaller than the ones calculated using the covariance matrix in SAS.

Figure 12 shows simulated bids using estimated parameters and Figure 13 shows the average bids for 30 such simulations. The individual simulation displays similar transition of bids over time compared to the actual bids in the experiment. The most prominent feature is the reduction in bid variance away from the LES bid function. The average of 30 simulations did a relatively poor job of capturing the magnitude of reduction in dispersion in asks.

Table 6 summarizes the estimated EWA learning parameters for sellers from experiment 2 using both the one- and two-segment models. The one-segment estimates indicate that initial attraction profile is approximately normal around a constant ask of \$20, with a standard deviation of 6.48. In contrast to the estimates for buyers, the decay parameters for cost and ask are similar in magnitude with more generalization for different strategies than for different costs. This is expected because the optimal bid function is very nonlinear for the buyers, resulting in little benefit to generalize the experience from different values. On the other hand, the bidding strategies for sellers are quite constant across costs, allowing for more useful generalization. The estimated value for ϕ is approximately one. The standard deviations of the one-segment model parameters are calculated from a bootstrap of 500 iterations with the results summarized in Table 7. Standard deviations obtained from this method are generally larger than the ones from the covariance matrix but of similar magnitude. A two segment model shows significant heterogeneity in the distribution of initial attractions as evidenced by the values of the log likelihood function, with a $\chi^2 = 39.17$ ($p < .005$). The two segments differ mostly in the level of the linear ask function. Figure 14 shows the simulation of sellers' asks using estimated parameters from the one-segment model. We see tightening of asks over the periods similar to the observed pattern in the experiment.

Direction Learning in Call Markets

Daniel et. al. looked at learning in this experiment using an adaptive model where a buyer increases or lower her bid in the form a one parameter family of exponential functions,

$$\text{Min}[V, h^*(1-\exp(-V/h-1))].$$

We can see from Figure 15 that this bid function is approximated by two linear segments, a full revelation strategy followed by an almost flat exponential function with the point of transition determined by the parameter of choice, h . Higher h produces bidding closer to values and a longer range of fully revealing bids. The right choice of parameter h fits the linear equilibrium quite well. In their adaptive model, successful trades lead buyers to lower their bids in the form of lower h and potentially profitable but unsuccessful bids lead to higher bids next period. This is in the same spirit as Selten and Buchta's learning direction theory, which only explains the direction of learning. In DSR's model, the strength of adjustment is proportional to either actual payoff when trade takes place or potential profit when a buyer misses a trade. Their learning model is described by,

$$B_t = \min[V_t, h_{t-1}(1 - \exp(-V_t/h_{t-1}))] \quad t = 1, 2, \dots$$

where

$$h_t = h_{t-1} \cdot [1 - w_t^+ (V_t - P_t)] \quad \text{if } B_t \geq A_t$$

$$h_t = h_{t-1} \cdot \max[1, 1 + w_t^- (V_t - A_t)] \quad \text{if } B_t < A_t$$

The strength of adjustment, w_t^+ and w_t^- are discounted by $w_t^+ = (1-d) w_{t-1}^+$ and $w_t^- = (1-d) w_{t-1}^-$, where d is a discount factor that depreciates the impact of outcomes.

They used a direction learning model similar to the one used for the buyers to explain the observed behavior of sellers in this market,

$$a_t = \max[C_p, c_{t-1} + m_{t-1} C_t], \quad t = 1, 2, \dots$$

where $m_t = \max(0, c_t/k - 1)$ is the slope of the (linear) ask function. The constant, k is chosen to minimize the RMSE of the model relative to observed choices and c_t is the parameter that determines the aggressiveness of the ask function by controlling both its intercept and slope. The principle of adjustment is analogous to the buyers' case, given by,

$$c_t = c_{t-1} [1 + w_{s,t}^+ (p_t - C_t)], \text{ when there is a trade and}$$

$$c_t = c_{t-1} \min(1, 1 - w_{s,t}^- (b_t - C_t)), \text{ when there is no trade,}$$

where $w_{s,t}^+ = (1-d) w_{s,t-1}^+$ and $w_{s,t}^- = (1-d) w_{s,t-1}^-$ are the incremental adjustment parameters that produce the optimal magnitude of adjustment with a constant discount of d . Thus after a trade, a seller will bid more aggressively by increasing both the intercept and slope of the linear function in direct proportion to her trading profit. When she missed a potentially

profitable trade, she will bid less aggressively next period by adjusting the linear function down using the difference between buyer's bid and her own cost.

They estimated individual parameters for each subject using the first 30 periods, then predicted the last 20 periods based on the estimated parameters. The average RMSEs for the first 30 and the last 20 periods are 9.31 and 8.73, respectively for the buyers and 13.39 and 7.54, respectively for the sellers in experiment 2. If we estimate the one-segment EWA learning model using the first 30 periods only, we get the following parameters for buyers and sellers,

Table 8: Estimated EWA learning parameters for bilateral call market using period 1-30.

Parameters	$N(0)$	$vmax$	α	β	γ	std	κ	ψ	ω	ϕ
Buyers	6.328	1.346	-.318	.393	.026	.191	.158	.012	.064	1.0
Sellers	3.418	36.761	15.515	2.107		12.941	.293	.036	.088	1.0

The RMSE of predicted choices for the first 30 and last 20 periods based on the estimates are 12.66 and 10.27 for the buyers and 17.24 and 12.03 for the sellers. The numbers are based on average choices from 30 simulations of the model since EWA is a stochastic choice model, whereas DSR's is a deterministic model. Clearly these numbers are not directly comparable because the EWA learning model was estimated for an "average" buyer and seller, but DSR's learning model is optimized for individual subjects. Using the averages of the individually estimated parameters given in DSR's paper produced RMSEs of 18.38 and 20.91 for the buyers and 19.24 and 10.86 for the sellers. There is not a clear preference for one model vs. another based on numerical comparison of predicted choices alone.

Now we look at the learning mechanisms driving both models. In bilateral call market, a good learning model needs to predict the direction and strength of adjustment each period. The direction learning model of DSR is particularly well suited to fit the bidding behavior of individual buyers in this institution because of the optimal two-segment bid function and the simple ordered strategy space. If a buyer makes a trade, the foregone payoffs will be less than the profit for all higher bids and higher than the realized profit for all lower bids down to the point of the actual ask. Thus the direction of modification, if any, should be to lower bids next period. On the other hand, if a buyer misses a trade but has a value higher than the seller's asking price, the direction is clearly to bid higher next period.

Once the direction of change is determined, DSR's model uses the data to determine the optimal discount rate to fit the observed speed of adjustment. The discount rate makes sure that the rate of adjustment diminishes continuously until it becomes negligible. We can see that DSR's learning model is a very parsimoniously way of fitting the observed adaptive behavior in this institution. However, it does miss some important qualitative characteristics that we observe in the data. In addition, it makes more sense to use foregone profit rather than actual profit to fit the strength of the adjustment when there is a trade because foregone profit was used when one misses a potentially profitable trade.

The same concern regarding buyers' adaptive behavior can be said about sellers in DSR's direction learning model. Namely, it is inconsistent to use actual payoff to motivate adaptive behavior when there is a trade but foregone payoff when there is no trade. EWA learning considers both actual and foregone payoffs for all choice adjustments. In addition, DSR's functional form is rather specific to the institution involved and the deterministic model does not allow for any stochastic property in subjects' choices. The convergence of choices is primarily driven by the discount parameter even as the actual payoffs increase with experience.

Conclusion

The application of EWA learning model to bilateral call markets extends a general learning model to market-like institutions with more complex strategies. Subjects in this market learn to partially generalize experience from one value/cost condition to another in response to the incentives from nonlinear optimal bid functions. The same learning model can be applied to other market institutions where subjects economize on learning by taking into consideration similarity between past experience and a new environment while still recognizing the difference in market incentives between them.

This application also demonstrated that “direction learning” is captured by EWA learning in a general way by taking into account foregone payoffs, without specifying a priori the direction of change. Direction learning was first proposed by Selten in the context of first price auctions, where the strategy space has a simple ordinal structure. When we use EWA learning in markets institutions with simple price strategies, the foregone payoffs from each choice produce the correct direction of adjustment without the ad hoc assumption on the direction of optimal strategies. In addition, EWA learning will also produce the strength of adjustment from foregone payoffs while direction learning model says nothing to that effect. The generality of EWA learning can further capture the confidence and extent with which subjects utilize the foregone payoffs to adjust their choices from period to period when the payoff information may be too complex or not fully realized.

Overall, EWA learning produces a generalized adaptive model of behavior that includes elements of reinforcement, belief-based and direction learning as special cases at some cost of complexity for the benefit of generality and logical consistency. It is a good foundation to build upon to extend our understanding of adaptive behaviors in more general games and market institutions. In future works, we should investigate the similarity parameters, ψ and ω , to better characterize their magnitude and significance in different market institutions

Figure 1: EWA cube

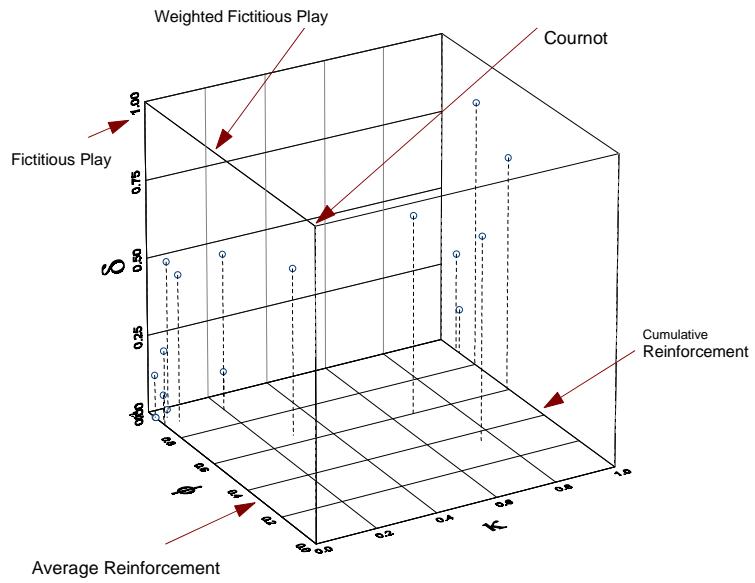


Figure 2: Time series of median choices in VCB's continental divide games.

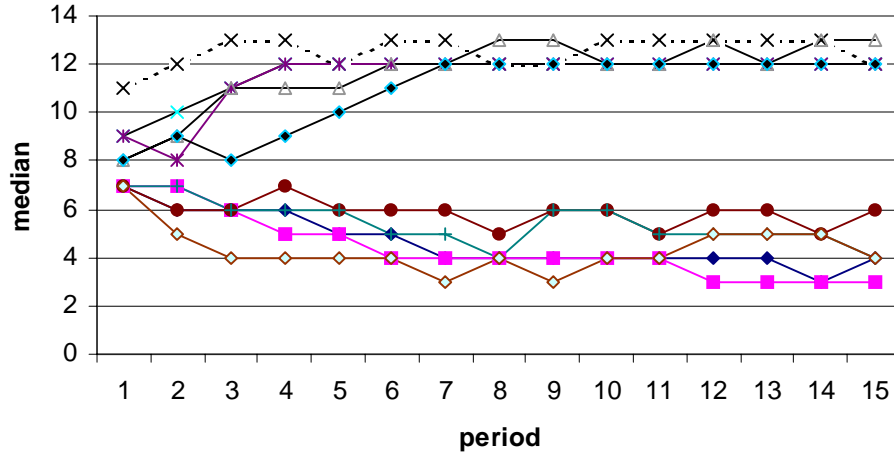


Figure 3: Distribution of choices in continental divide games.

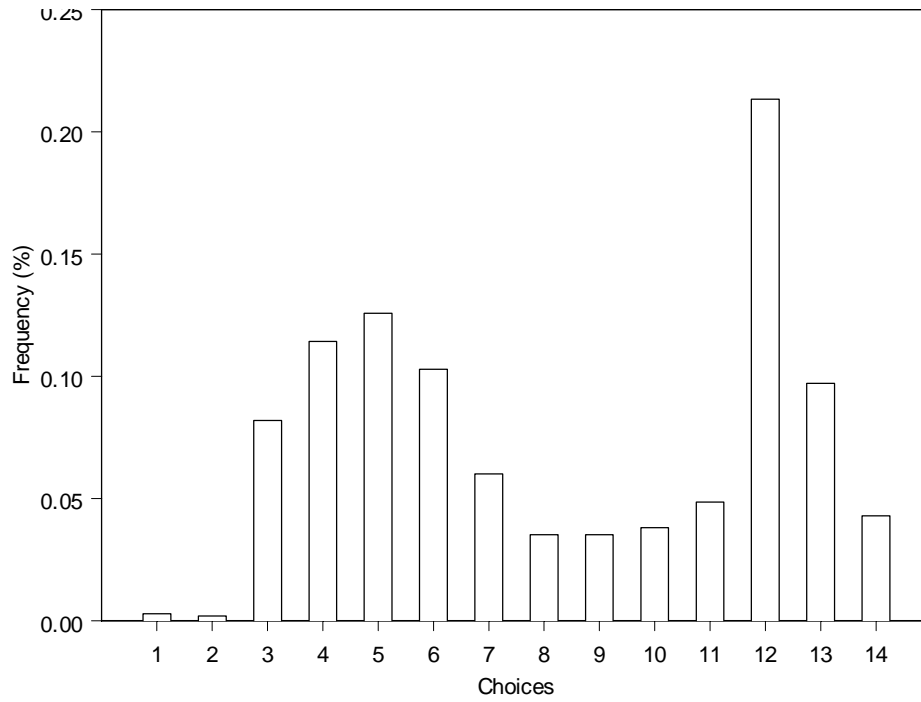


Figure 4: Comparison of simulated choices in continental divide games with learning models.

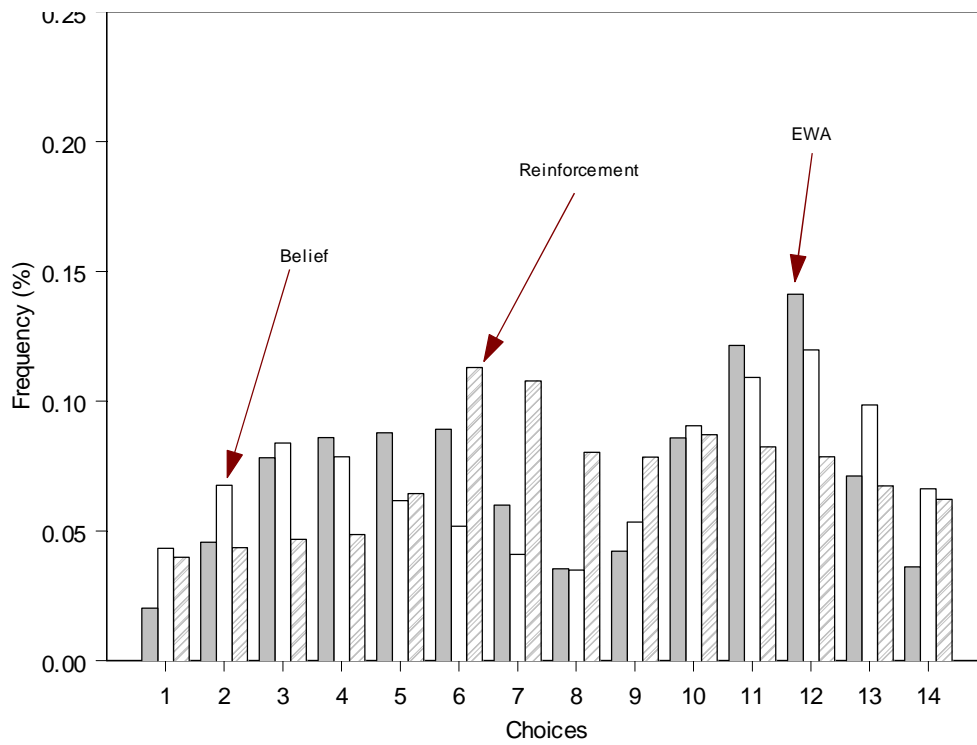


Figure 5: Sample time series of median choices in simulated continental divide games using EWA learning model.

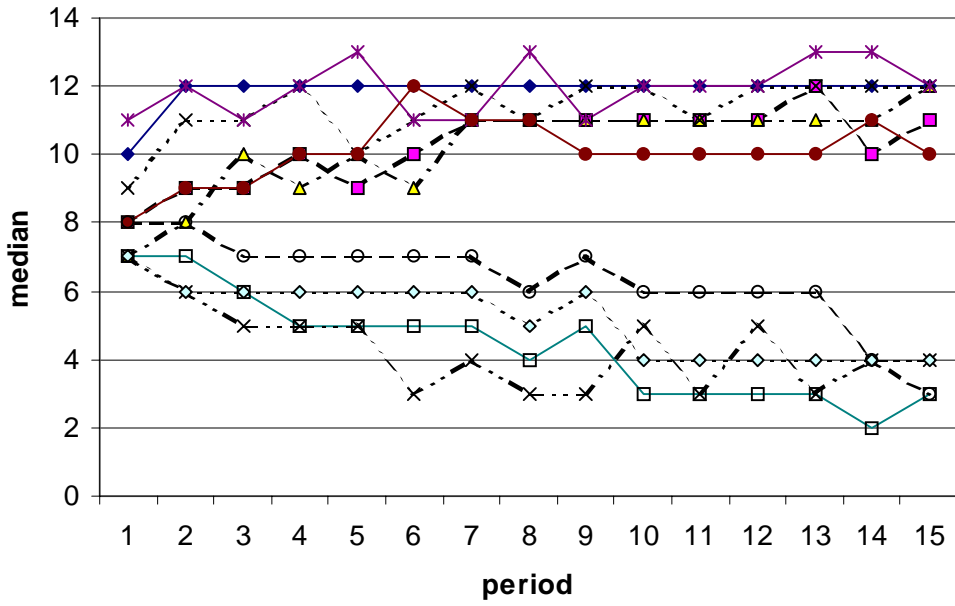


Figure 6: Sample time series of median choices in simulated continental divide games using reinforcement learning model.

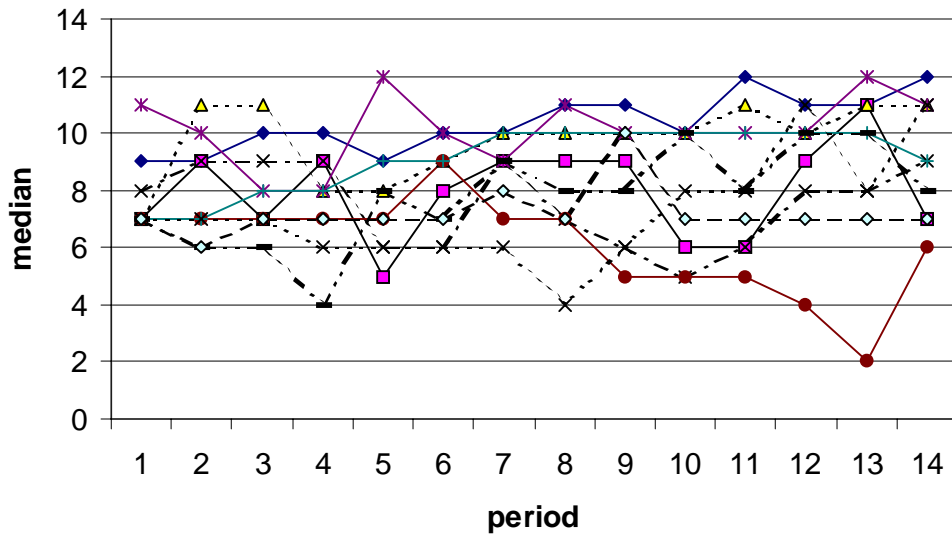


Figure 7: Sample time series of median choices in simulated continental divide games using belief learning model.

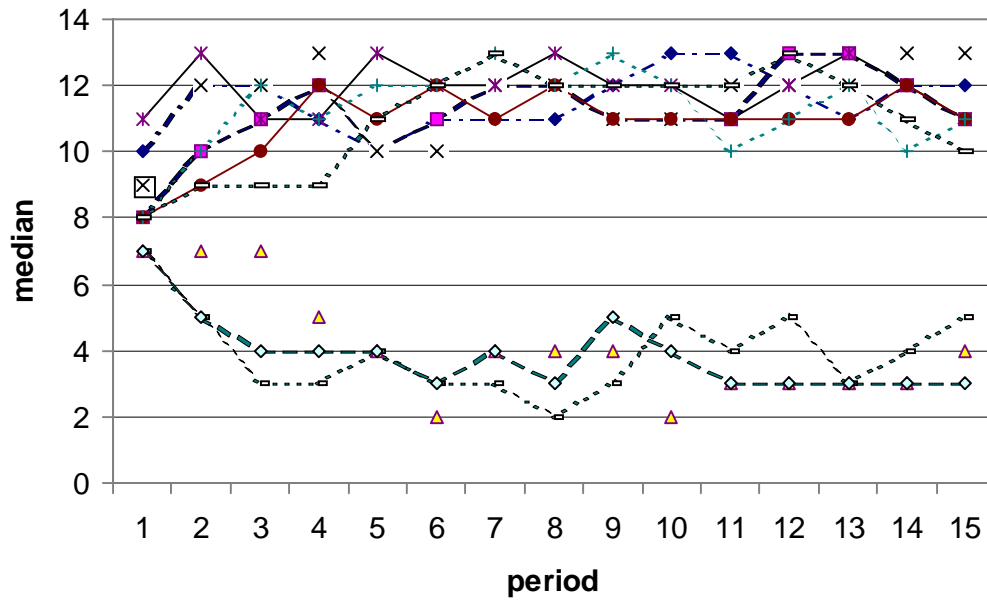


Figure 8. Comparison of actual bids vs. LES bid function

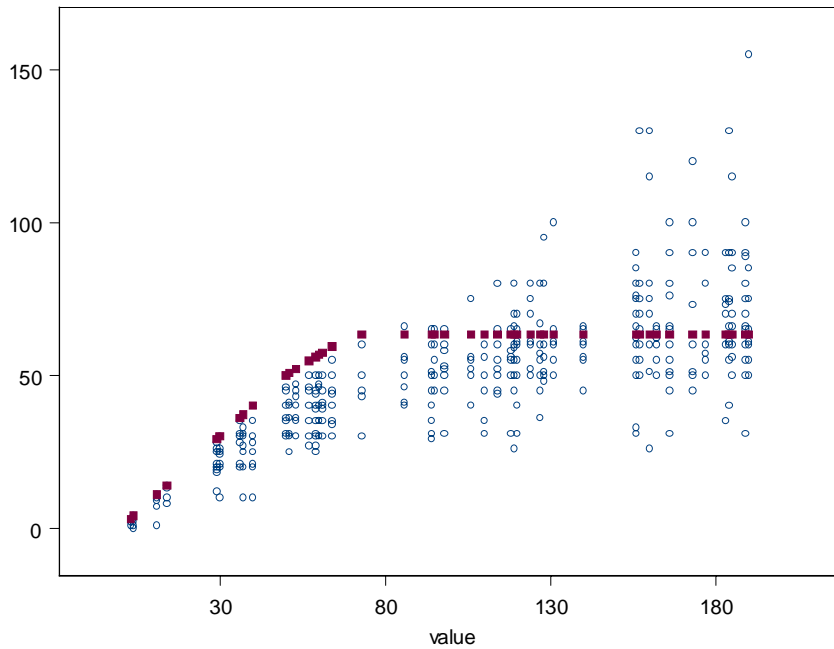


Figure 9 Distribution of buyers' bids over time for experiment 2.

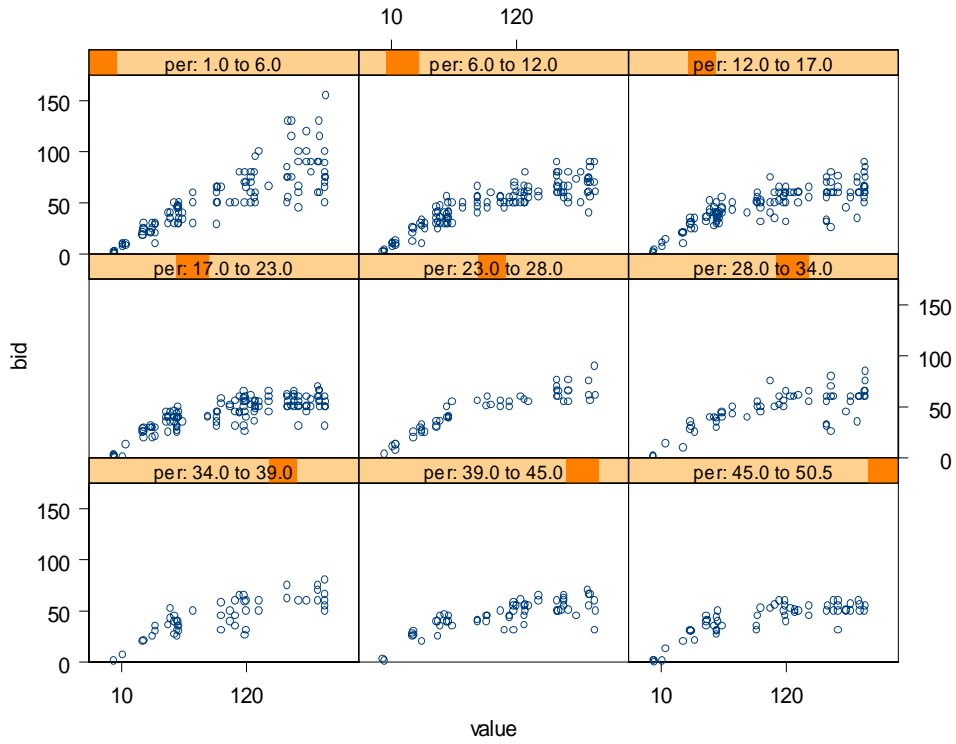


Figure 10. Distribution of sellers' asks over time for experiment 2

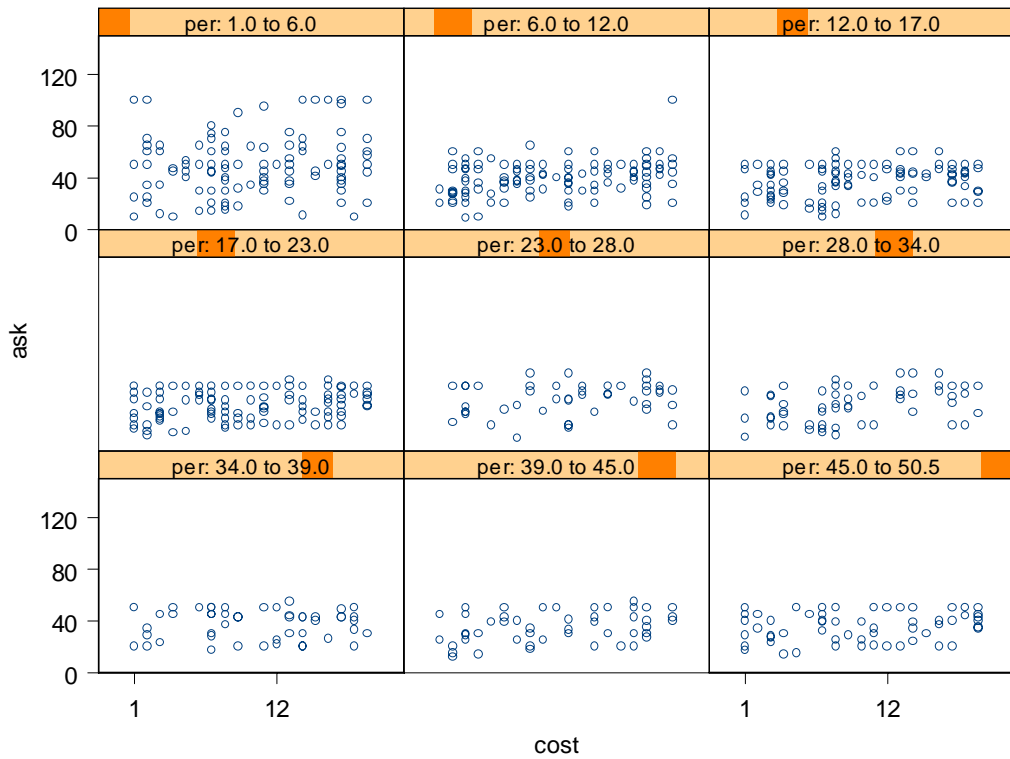


Figure 11. Distribution of generalization parameter, τ , for buyers with $v=100$, $s=30\%$ in experiment 2

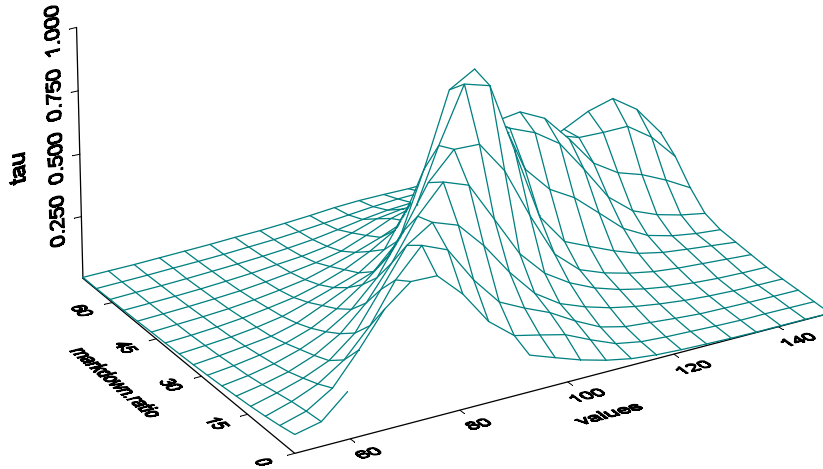


Figure 12. Simulated bids using EWA learning parameters for buyers in experiment 2.

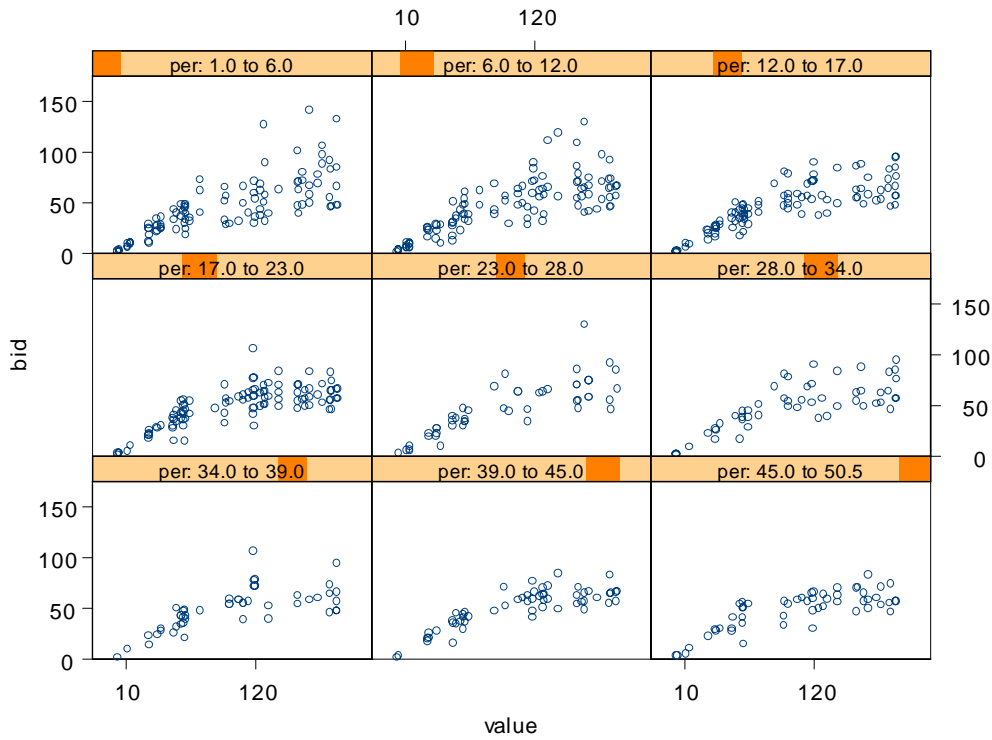


Figure 13: Average bids of 30 simulations using EWA learning parameters for experiment 2.

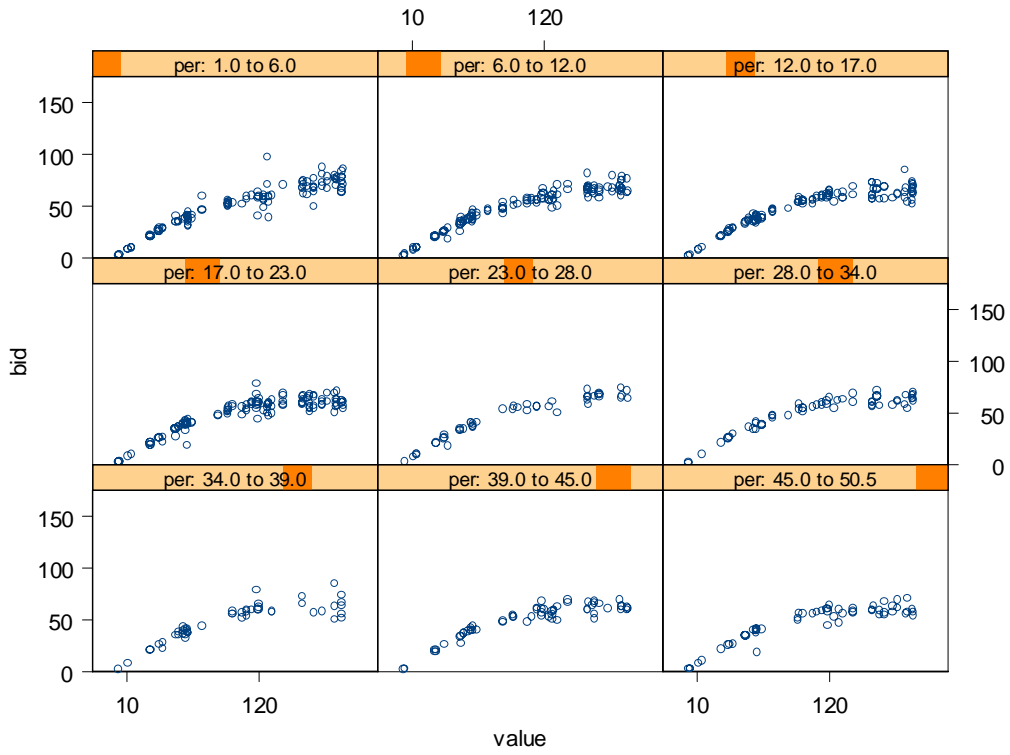


Figure 14: Simulated sellers' asks using estimated EWA learning parameters for experiment 2.

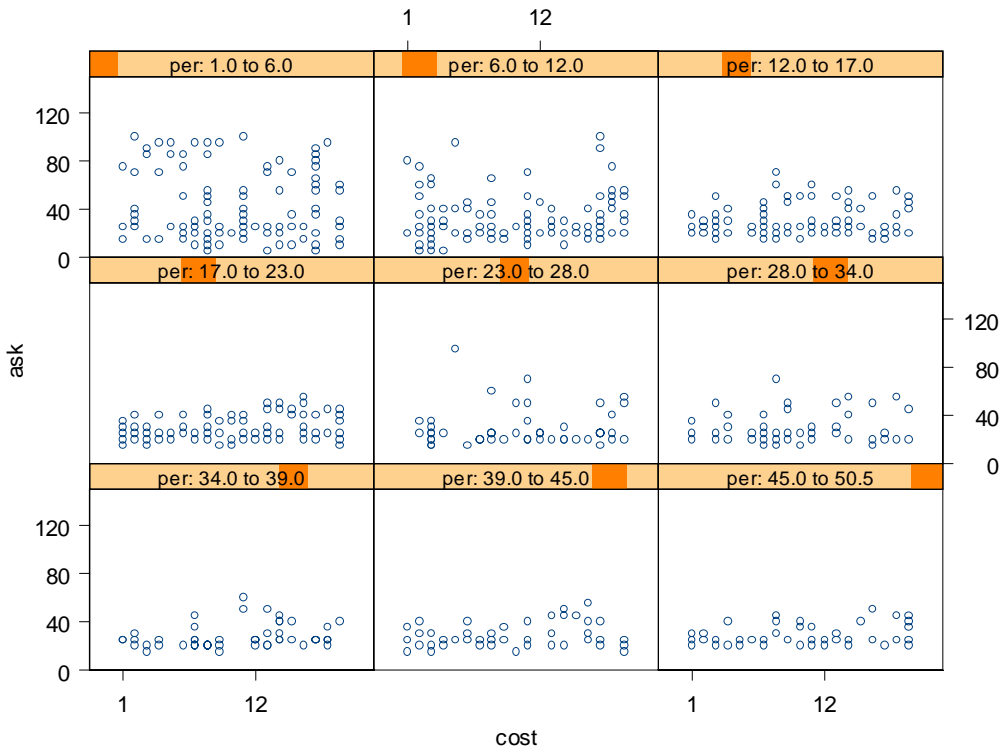


Figure 15. Comparison of LES bid function with DSR's one parameter exponential bid function with different values for parameter h.

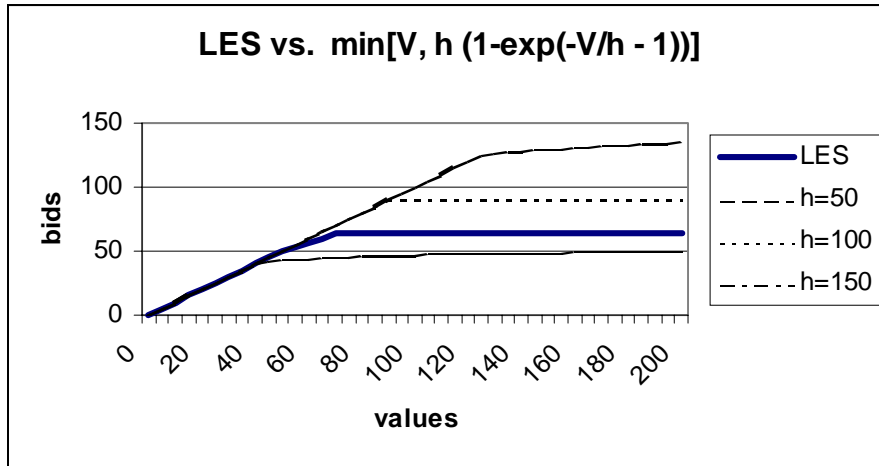


Table 1a: A summary of EWA parameter estimates in various games.

CITATION	GAME	EWA estimates (standard error)				Comments
		δ	ϕ	ρ	N(0)	
Camerer, Hsia and Ho (working)	Sealed bid mechanism	n.a.	1.00	0.91	11.2	ψ, ω & κ replace δ, ρ
Camerer, Ho and Weigelt (unpub'd)	“Continental divide”	0.750	0.611	0.001	0.762	
Hsia (unpublished)	Call markets	0.47 (0.32)	0.97 (0.01)	0.74 (0.06)	3.80 (0.88)	
Rapoport and Amaldoss (unpub'd)	Same function alliance – equal profit sharing	0.000	0.952	0.889	3.732	High reward
		0.000	0.985	0.986	1.773	Med. Reward
		0.070	0.892	0.867	0.110	Low reward
	Same function alliance – proportional sharing	0.000	0.988	0.996	15.176	High reward
		0.000	0.923	0.959	6.401	Med. Reward
		0.000	0.969	0.959	3.438	Low reward
	Parallel development of product – equal sharing	0.000	0.911	0.585	0.406	High reward
		0.168	0.903	0.934	4.910	Med. Reward
0.2112		0.883	0.645	0.236	Low reward	
Rapoport and Amaldoss (2000)	Patent race game – symmetric players	0.000	0.940	0.929	4.613	Low reward
		0.000	0.970	0.984	15.747	High reward
	Patent race game – asymmetric players	0.475	0.901	0.857	1.418	Strong player
		0.137	0.959	0.973	12.454	Weak player
Camerer and Anderson (1999)	Signaling games (game 3) (95% confidence interval)	0.69 (.47, 1.00)	1.02 (.99, 1.04)	1.00 (.98, 1.00)	32.9 32.8, 32.9	
	Signaling games (game 5) (95% confidence interval)	0.54 (.45, .63)	0.65 (.59, .71)	0.46 (.39, .54)	3.37 (3.4, 3.4)	
Camerer and Ho (1999)	Median-action coordination	0.853 (0.005)	0.800 (0.018)	0.000 (0.000)	0.647 (0.059)	
	4x4 Mixed-strategy games payoff = 5 rupees	0.000 (0.035)	1.04 (0.010)	0.961 (0.014)	19.63 (0.065)	
	4x4 Mixed-strategy games payoff = 10 rupees	0.73 (0.103)	1.005 (0.009)	0.946 (0.011)	18.391 (0.713)	
	6x6 Mixed-strategy games payoff = 5 rupees	0.413 (0.082)	0.986 (0.005)	0.935 (0.006)	15.276 (0.009)	
	6x6 Mixed-strategy games payoff = 10 rupees	0.547 (0.054)	0.991 (0.011)	0.926 (0.024)	9.937 (0.017)	
	p-beauty contest	0.232 (0.013)	1.330 (0.004)	0.941 (0.000)	16.815 (0.000)	
Camerer, Ho and Wang (1999)	Normal form centipede (odd player)	0.3190 (0.3157)	0.9092 (0.1414)	.0010	1.001 (2.2278)	Clairvoyance full update, κ
	Normal form centipede (even player)	0.239 (0.3218)	0.902 (0.1370)	.946	12.744 (4.9620)	Clairvoyance full update, κ
Chen (1999)	Cost allocation	.80~1.0	1 (fixed)	.1~.3	0 (fixed)	
Morgan and Sefton (1999)	“Unprofitable” games (baseline games)	0.084 (0.074)	0.925 (0.014)	0.916 (0.026)	3.632 (1.255)	
	“Unprofitable” games (upside games)	0.142 (0.064)	0.893 (0.012)	0.000 (0.004)	0.374 (0.343)	
Stahl (1999)	5x5 matrix games	.663 (.0211)	.337 (.0439)	.085 (.0821)	0.000 (0.000)	
Camerer and Ho (1998)	Weak-link coordination	0.652	0.582	0.198	2.187	

Table 1b: A summary of EWA forecast accuracy in various games.

CITATION	GAME	Model accuracy					Comments
		EWA	Choice reinforce - EWA	Belief - EWA	In / Out of sample	Fit technique	
Camerer, Hsia and Ho (working)	Sealed bid mechanism	1102	30.84	65.46	IN	-LL	ψ , ω & κ replace δ & ρ
Camerer, Ho and Weigelt (unpub'd)	“Continental divide”	1189.85	248.81	180.09	IN	-LL	
Hsia (unpublished)	Call markets	1915	0	403	IN	-LL	
Rapoport and Amaldoss (unpub'd)	Same function alliance – equal profit sharing	886.314	1.545	529.616	IN	-LL	High reward
		767.491	30.065	390.379	IN	-LL	Med. Reward
		1399.71	9.44	541.47	IN	-LL	Low reward
	Same function alliance – proportional sharing	910.777	36.432	812.953	IN	-LL	High reward
		1054.99	18.33	615.9	IN	-LL	Med. Reward
		1013.73	13.33	1095.61	IN	-LL	Low reward
	Parallel development of product – equal sharing	1194.15	0.11	566.31	IN	-LL	High reward
		1321.45	9.49	497.19	IN	-LL	Med. Reward
Rapoport and Amaldoss (2000)	Patent race game – symmetric players	3551.70	12.06	1097.69	IN	-LL	Low reward
		2908.08	20.21	725.94	IN	-LL	High reward
	Patent race game – asymmetric players	3031.54	89.06	706.77	IN	-LL	Strong player
		2835.51	15.65	610.98	IN	-LL	Weak player
Camerer and Anderson (in press)	Signaling games (game 3) (95% confidence interval)	72.16	6.48	10.08	OUT	-LL	
	Signaling games (game 5) (95% confidence interval)	139.52	14.08	23.68	OUT	-LL	
Camerer and Ho (1999)	Median-action coordination	41.05	39.22	72.85	OUT	-LL	
	4x4 Mixed-strategy games payoff = 5 rupees	326.38	9.12	-40.78	OUT	-LL	
	4x4 Mixed-strategy games payoff = 10 rupees	341.71	18.03	8.38	OUT	-LL	
	6x6 Mixed-strategy games payoff = 5 rupees	301.70	6.77	-5.42	OUT	-LL	
	6x6 Mixed-strategy games payoff = 10 rupees	362.26	13.68	8.92	OUT	-LL	
	p-beauty contest	2381.28	213.09	172.93	OUT	-LL	
Camerer, Ho and Wang (1999)	Normal form centipede (odd player)	1016.84	57.61	536.25	OUT	-LL	Clairvoyancefull update, κ
	Normal form centipede (even player)	951.30	46.42	604.7	OUT	-LL	Clairvoyancefull update, κ
Chen (1999)	Cost allocation	.73~.88	-.01~.07	n.a.	IN	MSD	
Morgan and Sefton (1999)	“Unprofitable” games (baseline games)	1729.47	0.59	n.a.	IN	-LL	
	“Unprofitable” games (upside games)	1906.45	16.15	n.a.	IN	-LL	
Stahl (1999)	5x5 matrix games	4803.73	64.68	n.a.	OUT	-LL	
Camerer and Ho (1998)	Weak-link coordination	358.058	29.105	438.546	IN	-LL	

Table 4: Summary of EWA learning parameter estimates for buyers in experiment 2

Parameters	Parameter estimates (standard errors from SAS)		
	One-segment model (SE)	Two-segment model	
N(0)	11.2 (67.4)	13.7 (6.8)	
Vmax	1.46 (14.2)	1.79 (1.4)	
α	0.04 (1.10)	0.09 (0.053)	-0.26 (0.106)
β	0.126 (0.73)	0.238 (0.125)	0.256 (0.065)
γ	0.072 (0.12)	0.048 (0.027)	0.052 (0.012)
Std. Dev.	0.219 (1.21)	0.210 (0.084)	0.208 (0.104)
Proportion		39.8%	60.2%
κ	0.09 (0.27)	0.07 (0.03)	
ψ	0.02 (0.05)	0.01 (0.01)	
ω	0.08 (0.12)	0.10 (0.03)	
ϕ	1.00 (constrained)	1.00	
-LL	1102.2	1071.1 ($\chi^2 = 62.2, p < .005$)	

Table 5: Summary of bootstrap parameter distribution for buyers in experiment 2 with λ and ϕ fixed at 1.0. (500 iterations)

	$N(0)$	VMAX	α	β	γ	STD	κ	ψ	ω
Average	3.987	1.101	-0.028	0.152	0.081	0.168	0.156	0.046	0.099
std dev.	0.104	0.473	0.194	0.134	0.106	0.062	0.066	0.040	0.056

Table 6: Summary of EWA learning parameters for sellers in experiment 2

Parameters	Parameter estimates (standard errors from SAS)		
	One-segment model	Two-segment model	
N(0)	3.29 (0.69)	3.27 (1.72)	
Vmax	19.24 (3.18)	14.43 (42.90)	
α	20.09 (2.59)	15.52 (34.90)	23.55 (8.66)
β	0.04 (0.19)	0.31 (2.20)	0.42 (0.22)
Std. Dev.	6.48 (1.12)	3.75 (17.85)	3.56 (7.39)
Proportion		78.3%	21.7%
κ	0.312 (0.022)	0.315 (0.071)	
ψ	0.091 (0.026)	0.089 (0.034)	
ω	0.059 (0.009)	0.048 (0.028)	
ϕ	1.011 (0.012)	1.014 (0.038)	
-LL	942.754	923.167	
		$\chi^2 = 39.17 (p < .005)$	

Table 7: Summary of bootstrap parameter distribution for sellers in experiment 2 with $\alpha=20$ and $\beta=0$. (500 iterations)

	$N(0)$	VMAX	STD	κ	ψ	ω	ϕ	LL
Average	3.155	19.348	7.575	0.313	0.087	0.061	1.013	-943.08
Std. Dev.	0.973	4.256	5.196	0.046	0.046	0.024	0.025	308.62

References

- Anderson, Christopher M. and Colin F. Camerer, Experience-Weighted Attraction in Sender-Receiver Signaling Games, *Economic Theory*, in press.
- Battalio, Raymond, Larry Samuelson and John Van Huyck, Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games, Working Paper, 1997.
- Camerer, Colin and Teck-Hua Ho, Experience Weighted Attraction Learning in Normal Form Games, *Econometrica*, 67, 1999, 827-873.
- Camerer, Colin, Teck H. Ho and Juin-Kuan Chong, Sophisticated EWA Learning and Strategic Teaching in Repeated Games, Caltech Working Paper, 1999.
- Camerer, Colin and Teck H. Ho. "Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity and time-variation," *Journal of Mathematical Psychology*, 42, 1998, 305-326.
- Cason, Timothy N. and Daniel Friedman, Learning in a Laboratory Market with Random Supply and Demand, *Experimental Economics*, 2(1), 1999, 77-98
- Colin, Camerer and Keith Weigelt, Experiential Tests of A Sequential Equilibrium Reputation Model, *Econometrica* Vol 56, no. 1, 1988, 1-36.
- Chen, Yan, Asynchrony and Learning in Serial and Average Cost Pricing Mechanisms: An Experimental Study, unpublished, 1999.
- Daniel, Terry E.; Seale, Darryl A.; and Amnon Rapoport. Strategic play and adaptive learning in the sealed-bid bargaining mechanism. *Journal of Mathematical Psychology*, 42, 1998, 133-166.
- Erev, Ido and Alvin E. Roth, Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria, *American Economic Review*, 88(4), 1998, 848-81.
- Friedman, Eric J., Learnability in a class on Non-Atomic Games arising on the Internet, working paper, 1998.
- Fudenberg, Drew and David K. Levine, *The Theory of Learning in Games*, MIT Press, 1998.
- Ho, Teck-Hua and Juin Kuan Chong. A parsimonious familiarity-based model of SKU choice. Wharton School working paper, 1999.
- Ho, Teck-Hua and Keith Weigelt, Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests", *The American Economic Review*, vol. 88, 1998, 947-969.
- Holt, Charles A., *Industrial Organization: A Survey of Laboratory Research*, in Kagel, John H. Roth, Alvin E., eds. *The Handbook of Experimental Economics*. Princeton: Princeton University Press, 1995, 349-443.
- Hsia, David W., *Learning in Single Call Markets*, manuscript, 1999
- Kaelbling, Leslie Pack, Michael L. Litman and Andrew W. Moore, Reinforcement Learning: A Survey, *Journal of Artificial Intelligence Research*, 4 1996, 237-285.

- Ledyard, John O. "The Design of Coordination Mechanisms and Organizational Computing." *Journal of Organizational Computing* 3, No. 1 (1993):121-134.
- Mookerjhee, D., and B. Sopher, Learning and Decision Costs in Experimental Constant-sum Games, *Games and Economic Behavior*, 19, 1997, 97-132.
- Mookerjhee, D., and B. Sopher, Learning Behavior in an Experimental Matching Pennies Game, *Games and Economic Behavior*, 7, 1994, 62-91.
- Morgan, John and Martin Sefton, An Experimental Investigation of Unprofitable Games, unpublished, 1998.
- Munro, Alistair. Rules of thumb in cycling coordination games. Unpublished, 1999.
- Radner, Roy and Andrew Schotter. The sealed-bid mechanism: An experimental study. *Journal of Economic Theory*, 48, 1989, 179-220.
- Rapoport, Amnon and Wilfred Amaldoss, Mixed Strategies and Iterative Elimination of Strongly Dominated Strategies: An Experimental Investigation of States of Knowledge, *Journal of Economic Behavior and Organization*, 2000.
- Rapoport, Amnon; Terry E. Daniel; and Darryl A. Seale. Reinforcement-based adaptive learning in asymmetric two-person bargaining with incomplete information. *Experimental Economics*, 1, 1998, 221-256.
- Rapoport, Amnon; and Mark A. Fuller. Bidding strategies in a bilateral monopoly with two-sided incomplete information. *Journal of Mathematical Psychology*, 39, 1995, 179-196.
- Roth, Alvin E. Bargaining experiments. In J. H. Kagel and A. E. Roth (eds.), *Handbook of Experimental Economics*. Princeton: Princeton University Press, 1995.
- Roth, Alvin E. and Ido Erev, Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term, *Games and Economic Behavior*, 8(1), 1995, 164-212.
- Salmon, Tim, Evidence for Learning to Learn Behavior in Normal Form Games, Working Paper, 2000.
- Salmon, Tim, An Evaluation of Econometric Models of Adaptive Learning, Working Paper 2000.
- Sarin, Rajiv and Farshid Vahid, Payoff Assessments without Probabilities: Incorporating "Similarity" Among Strategies, Texas A&M University Working Paper, 1997.
- Seale, Darryl A., Terry E. Daniel and Amnon Rapoport, The Information Advantage in Two-Person Bargaining with Incomplete Information, *Journal of Economic Behavior and Organization*, 2000.
- Selten, Reinhard and Joachim Butchta, Experimental Sealed Bid First Price Auction with Directly Observed Bid Functions, in *Games and Human Behavior: Essays in Honor of Amnon Rapoport*, ed by D. Budescu, I. Erev and R. Zwick, 1998.
- Schotter, Andrew; Snyder, Blaine; and Wei Zheng. Bargaining through agents: An experimental study. *Experimental Economics*, 2000.
- Stahl, Dale O. Evidence Based Rules and Learning in Symmetric Normal-Form Games, *International Journal of Game Theory*; 28, 1999, 111-30.

Van Huyck, John; Cook, Joseph; and Battalio, Raymond C. Adaptive behavior and coordination failure. *Journal of Economic Behavior and Organization*, 32, 1997, 483-503.

Van Huyck, John, Raymond C. Battalio and Frederick W. Rankin, Selection Dynamics and Adaptive Behavior Without Much Information, Texas A&M University Working Paper 1996.

Valley, Kathleen; Leigh Thompson; Robert Gibbons; and Max H. Bazerman. Using dyadic strategies to outperform equilibrium models of communication in bargaining games. Harvard Business School working paper 99-080, December 1998.

Weibull, Jorgen W., Evolution, Rationality and Equilibrium in Games, *European Economic Review*, 42(3-5), 1998, 641-49.

¹ Note well that while we assume players are reinforced by their monetary payoffs, the reinforcement function could easily be altered to account for loss-aversion, concave or convex utility for payoffs, or fairness.

² More precisely, if the special-case restriction is true, then the more general model will necessarily fit better in-sample, but will generally fit worse out-of-sample. Furthermore, criteria like the Bayesian information criterion, which penalizes a more complex theory for extra degrees of freedom, are carefully constructed so that even when sample sizes grow large, so that conventional test statistics appear to favor more complex theories, the extra-parameter penalty grows and there is no such bias.

³ Chen 's results are excluded because she did not estimate all parameters precisely. Results from the same game played at different stakes levels are collapsed together, and estimates are averaged across stakes levels. This reduces the 32 observations in Table 1 to 20 in Figure 1.