Identification of a General Learning Model on Experimental Game Data

Juergen Bracht
The Hebrew University of Jerusalem
Center for Rationality

Hidehiko Ichimura
University College London
Department of Economics
and
CeMMAP

December, 2001

Abstract

We discuss identification issues of general learning model and its special cases, reinforcement and belief learning (Camerer and Ho 1997, 1999). When the hypothesis of equilibrium play in Matching Pennies is not rejected, none of the parameters of the general model is identified. We show, nonparametrically, that this is indeed the case in Mookherjee and Sopher (1994) Matching Pennies data. Hence, the estimation results of Camerer and Ho (1997) are not reliable. There is also a case such that a subset of the parameters of the general model is not identified. We urge caution towards the interpretation of the estimation results in an important class of coordination games (Van Hyuck et al 1991). Our identification results enable us to map experimental games onto families of learning models. The results allow us to further investigate estimation issues of learning models.

1 Introduction

The experience weighted attraction (EWA) learning model proposed by Camerer and Ho (1999b) is a learning model in its own right however it can also be viewed as a convenient empirical framework to analyze learning behavior in experimental games. This is so because it encompasses reinforcement and belief learning models. From this perspective, identification of the model's parameters is an important issue. However, identification problems have not been carefully discussed in this context. We fill this gap in this paper.

We show that the EWA model's parameters are not always identified[1]. Identification of the parameters depend on the combinations of parameters that generate data.

We discuss two cases of non-identification and its empirical implications: First, consider repeated matrix games. Suppose that players use mixed strategies with probabilities which are equal and which stay constant over time. In this case, none of EWA model's parameters is identified. Further, suppose that those mixed strategies constitute a unique equilibrium. Then, the EWA model is incapable of informing us that subjects follow equilibrium play even if they did. Suppose we wish to examine learning models in such experimental games. This is inappropriate when we want to examine learning behavior without a priori restricting to a narrower class of models than the EWA model.

Second, suppose subjects' learning behavior is described by the reinforcement model. For this case, we show that some of the EWA parameters are not identified. Note that this result does not depend on the kind of game examined but depends only on the actual learning behavior. The result implies that not all aspects of reinforcement learning behavior can be inferred when subjects indeed follow the reinforcement learning behavior if we insist on examining the learning behavior within the class of models that includes the EWA model[2].

After establishing the identification results in section 2, we discuss its implications and review substantive findings in the literature that use the EWA learning model in light of our results. Section 4 concludes. All proofs are (will be) gathered in the Appendix.

2 Identification

2.1 EWA learning model

Camerer and Ho (1999) introduced the EWA learning model. We use the same notations as theirs but in order to discuss identification issues, we make parameters of the model explicit in this paper. Let the parameter of the EWA model, except for the adjustment speed parameter λ, be $\theta = \left(\phi, \delta, \rho, N(0), A^1(0), ..., A^S(0)\right)$. Players are indexed by i $(i = 1, 2, ..., n)$. The strategy space of the normal form game for player i is S_i. The number of elements in S_i is denoted as $\#S_i$ and it is assumed finite. An element s_i^j for $j = 1, ..., \#S_i$ in S_i denotes a strategy of player i. The scalar valued payoff function of player i is $\pi_i(s_i^{j_i}, s_{-i})$ where $s_{-i} = (s_1^{j_1}, s_2^{j_2}, ..., s_{i-1}^{j_{i-1}}, s_{i+1}^{j_{i+1}}, ..., s_n^{j_n})$ is the strategy combination of all players except player i. We denote the actual strategy chosen by player i in period t by $s_i(t)$ and his/her opponent's (or opponents') actual strategy vector by $s_{-i}(t)$. In this situation we denote ith player's payoff as $\pi_i(s_i(t), s_{-i}(t))$.

Two state variables, $N_i(t;\theta)$ and $A_i^s(t;\theta)$ $(s \in S_i)$ for each i in period t, are at the core of the EWA learning model. Here $N_i(t;\theta)$ is referred to as observation-equivalent of player i and controls the speed of learning. $A_i^s(t;\theta)$ is an indicator of player i's attraction to strategy s after period t has taken place. At each t, $\{A_i^s(t;\theta)\}_{s\in S_i}$ determines individual choice probabilities in the learning model. $N_i(t;\theta)$ affects the dynamics of $\{A_i^s(t;\theta)\}_{s\in S_i}$

[1] Hence, our results are contradicting Camerer and Ho (1999b) who claim that "It is easy to show algebraically that the parameters are identified" (p. 866).

[2] In recent work, Salmon (2001) has performed a similar task. He evaluates approaches to identify the learning model by simulation analysis. He focused on the question, if a particular model is in use, can these methods accurately identify it, for a given experimental game. He found that tests will typically suffer from significant type I and II errors.

but not the choice probability once $\{A_i^s(t;\theta)\}_{s\in S_i}$ is given. The EWA model specifies the initial observation equivalent, $N_i(0)$, and initial attractions, $A_i^s(0)$, how $N_i(t;\theta)$ and $A_i^s(t;\theta)$ are updated at each $t \geq 1$, and how attractions determine the strategy choice probabilities.

At the beginning of the first period of play observation equivalent of each player i is given by a parameter $N_i(0) = N(0)$ and initial attractions of each strategy are also left as some parameters i.e. for each $s \in S_i$, $A_i^s(0) = A^s(0)$ for $i = 1, 2, ..., n$.

After each period, attractions and observation equivalent of player i are updated. The updating rule of the observation equivalents is

$$N_i(t;\theta) = \rho \cdot N_i(t-1;\theta) + 1, \quad t \geq 1,$$

where ρ denotes the discount rate of observation equivalents. Note that $N_i(t;\theta)$ is the same for both players since we assume the same initial observation equivalent for all players and ρ is the same for all players, too. Hence, we drop the i subscript from $N_i(t;\theta)$ from now on.

Let $I\{A\} = 1$ if statement A is true and 0 otherwise. The updating rules of attractions are specified as

$$A_i^s(t;\theta) = \frac{\phi \cdot N(t-1;\theta) \cdot A_i^s(t-1;\theta) + [\delta + (1-\delta) \cdot I\{s_i(t) = s\}] \cdot \pi_i(s, s_{-i}(t))}{N(t;\theta)}, \quad t \geq 1,$$

where ϕ denotes a discount factor of attractions and δ denotes imagination.

Attractions determine how frequently players choose a particular strategy. The probability that player i chooses a strategy s in period $t+1$ is given by the following logit choice rule:

$$P_i^s(t+1;\lambda,\theta) = \frac{\exp(\lambda A_i^s(t;\theta))}{\sum_{s\in S_i} \exp(\lambda A_i^s(t;\theta))}, \quad t \geq 0,$$

where λ denotes sensitivity of players to attractions.

As Camerer and Ho (1999) discusses, the model is a reinforcement learning model if $\delta = 0$, $N(0) = 1$, and $\rho = 0$, and the model is a belief learning model if $\delta = 1$, $\phi = \rho$, and initial attractions are equal to expected payoffs given initial beliefs.

In the context of the reinforcement learning model, attractions are called propensities to choose strategies. This specification implies that individuals react only to the actual reward and that individuals do not adjust for the number of observation equivalents i.e. $N(t) = 1$ for all t. Let $Q_i^s(t)$ denote the propensity of strategy s of player i. The updating rule of propensities for $t \geq 1$ is

$$Q_i^s(t;\phi,\{A^s(0)\}_{s\in S_i}) = \phi \cdot Q_i^s(t-1;\phi,\{A^s(0)\}_{s\in S_i}) + I(s_i(t) = s) \cdot \pi_i(s, s_{-i}(t)).$$

In the context of the belief learning models, attractions are expected payoffs of strategies where expectation is based on past frequencies of opponents' chosen strategies possibly taking account of some discounting factor ϕ. Camerer and Ho (1999) shows that the belief learning models can be written as

$$E_i^s(t;\phi,\{A^s(0)\}_{s\in S_i}) = \frac{\phi \cdot N(t-1;\theta) \cdot E_i^s(t-1;\phi,\{A^s(0)\}_{s\in S_i}) + \pi(s_i^j, s_{-i}(t))}{N(t;\theta)}.$$

Note that expected payoffs are updated as a function of previous expected payoffs and current strategy payoffs.

Some learning models specify that

$$P_i^s(t+1;\theta) = \frac{A_i^s(t;\theta)}{\sum_{j\in S_i} A_i^j(t;\theta)}$$

and some other learning models, such as the EWA model, specify

$$P_i^s(t+1;\lambda,\theta) = \frac{\exp[\lambda \cdot A_i^s(t;\theta)]}{\sum_{j\in S_i} \exp\left[\lambda \cdot A_i^j(t;\theta)\right]},$$

where λ is the adjustment speed parameter.

While the first specification is invariant to the proportional change in the attractions the second is invariant to the additive change in the attractions. When the ith player arrives at the same proportion of attractions in earlier period and in later period, the first formulation predicts that the player chooses the strategies with the same probabilities but the second formulation predicts that the player chooses the strategy that has the highest attraction more at a later period than at an earlier period. This is because $\sum_{j\in S_i} A_i^j(t;\theta)$ is monotonically increasing. Analogously, when the ith player arrives at the same differences of attractions in earlier period and in later period, the second formulation predicts that the player chooses the strategies with the same probabilities but the first formulation predicts that the player chooses the strategy that has the highest attraction less at a later period than at an earlier period.

In this paper we examine identification of the parameters in the EWA model in its original form.

As we discussed the EWA learning model specifies the choice probabilities over the strategy space given the past choices and outcome of all players as a nonlinear (stochastic) difference equations in $\{A_i^s(t;\theta)\}_{s=1}^{S_i}$ and $N(t;\theta)$. Since the equations $\{A_i^s(t;\theta)\cdot N(t;\theta)\}_{s=1}^{S_i}$ is a linear (stochastic) difference equations, we can solve the problem explicitly. This result enables us to obtain the precise identification results.

Theorem 1 *The choice probability distribution over S_i in the first period is*

$$P_i^s(1;\lambda,\theta) = \frac{\exp(\lambda A^s(0))}{\sum_{j\in S_i}\exp(\lambda A^j(0))}$$

The choice probability distribution over S_i in period $t+1\geq 2$ in terms of past choices and payoffs in EWA learning model is

$$P_i^s(t+1;\lambda,\theta) = \frac{\exp(\lambda A_i^s(t;\theta))}{\sum_{j\in S_i}\exp(\lambda A_i^j(t;\theta))}$$

where

$$A_i^s(t;\theta) = \frac{\phi^t\cdot A^s(0)\cdot N(0) + \sum_{\tau=0}^{t-1}\phi^\tau\cdot[\delta+(1-\delta)\cdot I\{s_i(t-\tau)=s\}]\cdot\pi_i(s,s_{-i}(t-\tau))}{1+\rho+\cdots+\rho^{t-1}+\rho^t N(0)}.$$

The proof is in the appendix.

2.2 Identification results

The following lemma is elementary but useful for our purpose.

Lemma 2 $1+\rho+\cdots+\rho^{t-1}+\rho^t N(0) = 1/(1-\rho)$ *when* $N(0)=1/(1-\rho)$.

It implies that if ρ and ϕ are close to 1 and $N(0)=1/(1-\rho)$, $A_i^s(t;\theta)$ will be approximately constant regardless of history. Note that this holds regardless of δ values and λ values. Another case is when $N(0)(1-\rho)$ and $1+\rho+\cdots+\rho^{t-1}+\rho^t N(0)$ both diverge and that $\rho=\phi$, then $A_i^s(t;\theta)$ will be approximately constant. Hence we have

Theorem 3 *For any numbers p^s, $0<p^s<1$ for $s\in S_i$ with $\sum_{s\in S_i} p^s = 1$ there are non-unique solutions to the following set of equations: $P_i^s(t+1;\lambda,\theta)=p^s$.*

For clarity we state the definition of identification first. Suppose (λ_0,θ_0) underlies the joint distribution of $\{s_i(t)\}$ for $i=1,...,n$ and $t=1,...,T$ that generated the data. Let Γ be a function which maps (λ,θ) to an element in a finite dimensional real space.

Definition 4 *$\Gamma(\lambda_0, \theta_0)$ is identified from $\Gamma(\lambda, \theta)$ if and only if the same joint distribution of $\{s_i(t)\}$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T$ implies $\Gamma(\lambda_0, \theta_0) = \Gamma(\lambda, \theta)$. That is*

$$P_i^s(t; \lambda, \theta) = P_i^s(t; \lambda_0, \theta_0)$$

with probability 1 for each $t = 1, \ldots, T$ implies $\Gamma(\lambda_0, \theta_0) = \Gamma(\lambda, \theta)$.

If Γ can be taken to be an identity map, we say (λ_0, θ_0) is identified.

The next theorem clarifies the identification conditions of the EWA model.

Define for $s = 2, \ldots, S$ and $t = 1, \ldots, T$

$$\begin{aligned}
\Delta\pi_{it}^s &= I\{s_i(t) = s\} \cdot \pi_i(s, s_{-i}(t)) - I\{s_i(t) = 1\} \cdot \pi_i(1, s_{-i}(t)) \\
\Delta\tilde{\pi}_{it}^s &= I\{s_i(t) \neq s\} \cdot \pi_i(s, s_{-i}(t)) - I\{s_i(t) \neq 1\} \cdot \pi_i(1, s_{-i}(t))
\end{aligned}$$

It turns out variables I_{it}^s, $\Delta\pi_{it}^s$, and $\Delta\tilde{\pi}_{it}^s$ play role of regressors. The following lemma shows that the regressors will not be degenerate so long as the EWA parameters are finite and the payoff matrix satisfies certain regularity conditions.

Lemma 5 *Suppose the EWA parameters and payoffs are all finite and for any $s \in S_i$ for each $i = 1, \ldots, n$, there exists at least one strategy combination by others which make the payoff to be different, then any linear combination of $I_{i1}^2, \ldots, I_{i1}^S, \Delta\pi_{i1}^{s_i(1)}, \Delta\tilde{\pi}_{i1}^{s_i(1)}, \ldots, I_{it}^2, \ldots, I_{it}^S, \Delta\pi_{it}^{s_i(t)}, \Delta\tilde{\pi}_{it}^{s_i(t)}$ has a positive variance.*

We will call the conditions stated in the lemma as a regular case.

Theorem 6 *Consider a regular EWA model. If for $t = 2$ and for each $s = 2, \ldots, S$, then*

$$\begin{gathered}
\lambda_0 \left[A_0^s(0) - A_0^1(0) \right] \text{ for } s = 2, \ldots, S, \\
\frac{\lambda_0 \cdot \phi_0 \left[A_0^s(0) - A_0^1(0) \right] \cdot N_0(0)}{1 + \rho_0 N_0(0)} \text{ for } s = 2, \ldots, S, \\
\frac{\lambda_0 \cdot \delta_0}{1 + \rho_0 N_0(0)}, \\
\frac{\lambda_0}{1 + \rho_0 N_0(0)}
\end{gathered}$$

are identified. Furthermore, δ_0 is identified if and only if $\lambda_0 \neq 0$ and $|\rho_0 N_0(0)| < \infty$.

For $t \geq 3$

$$\begin{gathered}
\lambda_0 \left[A_0^s(0) - A_0^1(0) \right] \text{ for } s = 2, \ldots, S, \\
\frac{\lambda_0 \cdot \phi_0^t \left[A_0^s(0) - A_0^1(0) \right] \cdot N_0(0)}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)} \text{ for } s = 2, \ldots, S, \\
\frac{\lambda_0 \cdot \delta_0 \cdot \phi^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}, \\
\frac{\lambda_0 \cdot \phi_0^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}.
\end{gathered}$$

are identified. In this case δ_0 and ϕ_0 are identified if and only if $\lambda_0 \neq 0$ and $|\rho_0 N_0(0)| < \infty$. In addition all the parameters are identified after normalizing $A_0^1(0) = 0$, if and only if $\lambda_0 \neq 0$ and $|\rho_0 N_0(0)| < \infty$ and

$$\frac{N_0(0)}{1 + \rho_0 N_0(0)} \neq 1 \text{ or equivalently when } \rho_0 = 1 \text{ or } 0 \leq \rho_0 < 1 \text{ and } N_0 \neq \frac{1}{1 - \rho_0}.$$

Note that if $N_0(1-\rho_0)=1$, then $\lambda_0(1-\rho_0)$ is identified but not more.

Proof of Theorem 2:

We first prove the following lemma:

Lemma 7 *Suppose λ, ϕ, δ, $\rho > 0$, $N(0) > 0$ and $A^s(0)$ for each $s \in S_i$ are all finite and payoff is also finite. Then for each $i = 1, ..., n$, $s \in S_i$ and $t = 1, ..., T < \infty$, $0 < P_i^s(t; \lambda, \theta) < 1$ in the EWA learning model.*

Proof: Since $N(t) = 1 + \rho + \cdots \rho^{t-1} + \rho^t N(0)$, $N(t)$ is greater than 1 and bounded for all finite t when ρ and $N(0)$ are nonnegative and bounded. Since

$$P_i^s(1; \lambda, \theta) = \frac{\exp(\lambda A^s(0))}{\sum_{j \in S_i} \exp(\lambda A^j(0))},$$

the result holds as $-\infty < \lambda A^s(0) < \infty$ for each $s \in S_i$ and $i = 1, ..., n$ and $t = 1$. Suppose $-\infty < \lambda A_i^s(t; \lambda, \theta) < \infty$ holds for a $t = 0, ..., T$. We show that the result also holds for $t+1$. Since

$$P_i^s(t+1; \lambda, \theta) = \frac{\exp(\lambda A_i^s(t; \theta))}{\sum_{j \in S_i} \exp(\lambda A_i^j(t; \theta))}$$

the result implies the claim of the lemma. To see the induction argument holds, note that

$$A_i^s(t+1; \lambda, \theta) = \frac{\phi \cdot N(t) \cdot A_i^s(t; \lambda, \theta) + [\delta + (1-\delta) \cdot I\{s_i(t) = s\}] \cdot \pi_i(s, s_{-i}(t))}{N(t+1)}, \quad t \geq 1,$$

when $N(t)$ is greater than 1 and bounded for any finite t, and when ϕ, δ, $\pi_i(s, s_{-i}(t))$, and $A_i^s(t; \lambda, \theta)$ are all finite, $A_i^s(t+1; \lambda, \theta)$ is also finite. ¤

Note that for each period there are $\#S_i$ equalities but if any $\#S_i - 1$ equalities hold the remaining equality must hold because the equations sum to one. Hence we can examine $\#S_i - 1$ equalities without a loss of generality. We choose the $\#S_i - 1$ equalities we get after taking the ratio with respect to the first choice for each period. This result in the following set of equalities: for $s = 2, ..., \#S_i$

$$\frac{\exp(\lambda A^s(0))}{\exp(\lambda A^1(0))} = \frac{\exp(\lambda_0 A_0^s(0))}{\exp(\lambda_0 A_0^1(0))}$$
$$\frac{\exp(\lambda A_i^s(1; \theta))}{\exp(\lambda A_i^1(1; \theta))} = \frac{\exp(\lambda_0 A_i^s(1; \theta_0))}{\exp(\lambda_0 A_i^1(1; \theta_0))}.$$

Taking log and using Theorem 1 and rearranging, we obtain, for $s = 2, ..., \#S_i$

$$\lambda \Delta A^s = \lambda_0 \Delta A_0^s$$
$$\frac{\lambda \cdot \phi \cdot \Delta A^s(0) \cdot N(0) + \lambda \cdot \Delta\pi_i^s + \lambda \cdot \delta \cdot \Delta\tilde{\pi}_i^s}{1 + \rho N(0)} = \frac{\lambda_0 \cdot \phi_0 \cdot \Delta A_0^s(0) \cdot N_0(0) + \lambda_0 \cdot \Delta\pi_i^s + \lambda_0 \cdot \delta_0 \cdot \Delta\tilde{\pi}_i^s}{1 + \rho_0 N_0(0)},$$

where

$$\begin{aligned}
\Delta A^s &= A^s(0) - A^1(0) \\
\Delta A_0^s &= A_0^s(0) - A_0^1(0) \\
\Delta\pi_{it}^s &= I\{s_i(t) = s\} \cdot \pi_i(s, s_{-i}(t)) - I\{s_i(t) = 1\} \cdot \pi_i(1, s_{-i}(t)) \\
\Delta\tilde{\pi}_{it}^s &= I\{s_i(t) \neq s\} \cdot \pi_i(s, s_{-i}(t)) - I\{s_i(t) \neq 1\} \cdot \pi_i(1, s_{-i}(t)).
\end{aligned}$$

When

Writing this out more explicitly we have, for $s = 1, ..., S$

$$\frac{\exp(\lambda A^s(0))}{\sum_{j\in\mathcal{S}} \exp(\lambda A^j(0))} = \frac{\exp(\lambda_0 A_0^s(0))}{\sum_{j\in\mathcal{S}} \exp(\lambda_0 A_0^j(0))}$$
$$\frac{\exp(\lambda A_i^s(1;\theta))}{\sum_{j\in\mathcal{S}} \exp(\lambda A_i^j(1;\theta))} = \frac{\exp(\lambda_0 A_i^s(1;\theta_0))}{\sum_{j\in\mathcal{S}} \exp(\lambda_0 A_i^j(1;\theta_0))}$$
$$\vdots$$
$$\frac{\exp(\lambda A_i^s(T;\theta))}{\sum_{j\in\mathcal{S}} \exp(\lambda A_i^j(T;\theta))} = \frac{\exp(\lambda_0 A_i^s(T;\theta_0))}{\sum_{j\in\mathcal{S}} \exp(\lambda_0 A_i^j(T;\theta_0))}.$$

Note that for each period there are S equalities but if any $S-1$ equalities hold the remaining equality must hold because the equations sum to one. Hence we can examine $S-1$ equalities without a loss of generality. We choose the $S-1$ equalities we get after taking the ratio with respect to the first choice for each period. This result in the following set of equalities: for $s = 1, ..., S$

$$\frac{\exp(\lambda A^s(0))}{\exp(\lambda A^1(0))} = \frac{\exp(\lambda_0 A_0^s(0))}{\exp(\lambda_0 A_0^1(0))}$$
$$\frac{\exp(\lambda A_i^s(1;\theta))}{\exp(\lambda A_i^1(1;\theta))} = \frac{\exp(\lambda_0 A_i^s(1;\theta_0))}{\exp(\lambda_0 A_i^1(1;\theta_0))}$$
$$\vdots$$
$$\frac{\exp(\lambda A_i^s(T;\theta))}{\exp(\lambda A_i^1(T;\theta))} = \frac{\exp(\lambda_0 A_i^s(T;\theta_0))}{\exp(\lambda_0 A_i^1(T;\theta_0))}$$

Define for $s = 2, ..., S$ and $t = 1, ..., T$

$$\begin{aligned}
\Delta\pi_i(s, s_{-i}(t)) &= \pi_i(s, s_{-i}(t)) - \pi_i(1, s_{-i}(t)) \\
Z_i(s, s_{-i}(t)) &= I\{s_i(t) = s\} \cdot \pi_i(s, s_{-i}(t)) - I\{s_i(t) = 1\} \cdot \pi_i(1, s_{-i}(t)) \\
X_i(s, s_{-i}(t)) &= \Delta\pi_i(s, s_{-i}(t)) - Z_i(s, s_{-i}(t)) = I\{s_i(t) \neq s\} \cdot \Delta\pi_i(s, s_{-i}(t))
\end{aligned}$$

Using the expression for the attraction for the initial period we have, for $s = 2, ..., S$,

$$\lambda \left[A^s(0) - A^1(0)\right] = \lambda_0 \left[A_0^s(0) - A_0^1(0)\right].$$

Also for the first period we have, for $s = 2, ..., S$

$$\lambda \cdot \left[\frac{\phi \left[A^s(0) - A^1(0)\right] \cdot N(0) + \delta \cdot X_i(s, s_{-i}(1)) + Z_i(s, s_{-i}(1))}{1 + \rho N(0)}\right]$$
$$= \lambda_0 \cdot \left[\frac{\phi_0 \left[A_0^s(0) - A_0^1(0)\right] \cdot N_0(0) + \delta_0 \cdot X_i(s, s_{-i}(1)) + Z_i(s, s_{-i}(1))}{1 + \rho_0 N_0(0)}\right].$$

Analogously, in general we have, for $t = 1, ..., T$

$$\lambda \cdot \left[\frac{\phi^t \cdot \left[A^s(0) - A^1(0)\right] \cdot N(0) + \sum_{\tau=0}^{t-1} \left[\phi^\tau \cdot \delta \cdot X_i(s, s_{-i}(t-\tau)) + \phi^\tau \cdot Z_i(s, s_{-i}(t-\tau))\right]}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)}\right]$$
$$= \lambda_0 \cdot \left[\frac{\phi_0^t \cdot \left[A_0^s(0) - A_0^1(0)\right] \cdot N_0(0) + \sum_{\tau=0}^{t-1} \left[\phi_0^\tau \cdot \delta_0 \cdot X_i(s_i(t-\tau), s_{-i}(t-\tau)) + \phi_0^t \cdot Z_i(s_i(t-\tau), s_{-i}(t-\tau))\right]}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}\right].$$

>From the first period equality, if for each $s = 2, ..., S$

$$\left(I\left\{s_i(1) = s\right\}, X_i\left(s, s_{-i}(1)\right), Z_i\left(s, s_{-i}(1)\right)\right)$$

has a nondegenerate distribution in the area with positive probability, we have

$$\begin{aligned}
\frac{\lambda \cdot \phi \left[A^s(0) - A^1(0)\right] \cdot N(0)}{1 + \rho N(0)} &= \frac{\lambda_0 \cdot \phi_0 \left[A_0^s(0) - A_0^1(0)\right] \cdot N_0(0)}{1 + \rho_0 N_0(0)}, \\
\frac{\lambda \cdot \delta}{1 + \rho N(0)} &= \frac{\lambda_0 \cdot \delta_0}{1 + \rho_0 N_0(0)}, \\
\frac{\lambda}{1 + \rho N(0)} &= \frac{\lambda_0}{1 + \rho_0 N_0(0)}.
\end{aligned}$$

Analogously, if for some t

$$\left(I\left\{s_i(1) = s\right\}, X_i\left(s_i(1), s_{-i}(1)\right), Z_i\left(s_i(1), s_{-i}(1)\right), ..., I\left\{s_i(t) = s\right\}, X_i\left(s_i(t), s_{-i}(t)\right), Z_i\left(s_i(t), s_{-i}(t)\right)\right)$$

has a nondegenerate distribution in the area with positive probability, we have for $\tau = 0, ..., t-1$, $t = 1, ..., T$

$$\begin{aligned}
\frac{\lambda \cdot \phi^t \left[A^s(0) - A^1(0)\right] \cdot N(0)}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} &= \frac{\lambda_0 \cdot \phi_0^t \left[A_0^s(0) - A_0^1(0)\right] \cdot N_0(0)}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}, \\
\frac{\lambda \cdot \delta \cdot \phi^\tau}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} &= \frac{\lambda_0 \cdot \delta_0 \cdot \phi^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}, \\
\frac{\lambda \cdot \phi^\tau}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} &= \frac{\lambda_0 \cdot \phi_0^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}.
\end{aligned}$$

It is important to recognize that the nondegeneracy of the regressors condition itself depends on the underlying behavioral model.

Since shifting the attraction times λ by a constant amount does not change the probabilities we impose the restriction that $A^1(0) = 0$. With this normalization we see that

$$\begin{aligned}
\frac{\exp(\lambda A^s(0))}{\exp(\lambda A^1(0))} &= \frac{\exp(\lambda_0 A_0^s(0))}{\exp(\lambda_0 A_0^1(0))} \\
\frac{\exp(\lambda A_i^s(1;\theta))}{\exp(\lambda A_i^1(1;\theta))} &= \frac{\exp(\lambda_0 A_i^s(1;\theta_0))}{\exp(\lambda_0 A_i^1(1;\theta_0))} \\
&\vdots \\
\frac{\exp(\lambda A_i^s(T;\theta))}{\exp(\lambda A_i^1(T;\theta))} &= \frac{\exp(\lambda_0 A_i^s(T;\theta_0))}{\exp(\lambda_0 A_i^1(T;\theta_0))}
\end{aligned}$$

Analogously the probability distribution of $s_i(t+1)$ given $\{s_k(\tau)\}_{k=1}^n$ for $\tau = 1, ..., t$ is assumed random over i with $s = 1, ..., S$

$$P_i^s(t+1;\lambda,\theta) = \frac{\exp\left\{\frac{\phi \cdot A^s(0) \cdot N(0) + \sum_{\tau=0}^{t-1} \phi^\tau \cdot [\delta + (1-\delta) \cdot I\{s_i(t-\tau) = s\}] \cdot \pi_i(s, s_{-i}(t-\tau))}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)}\right\}}{\sum_{j=1}^S \exp\left\{\frac{\phi \cdot A^j(0) \cdot N(0) + \sum_{\tau=0}^{t-1} \phi^\tau \cdot [\delta + (1-\delta) \cdot I\{s_i(t-\tau) = j\}] \cdot \pi_i(j, s_{-i}(t-\tau))}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)}\right\}}.$$

the following parameters are identified: for $s = 2, ..., S$, and for $\tau = 0, ..., t-1$, $t = 1, ..., T$,

$$
\begin{aligned}
\lambda \left[A^s(0) - A^1(0)\right] &= \lambda_0 \left[A_0^s(0) - A_0^1(0)\right], \\
\frac{\lambda \cdot \phi^t \left[A^s(0) - A^1(0)\right] \cdot N(0)}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} &= \frac{\lambda_0 \cdot \phi_0^t \left[A_0^s(0) - A_0^1(0)\right] \cdot N_0(0)}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}, \\
\frac{\lambda \cdot \delta \cdot \phi^\tau}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} &= \frac{\lambda_0 \cdot \delta_0 \cdot \phi^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}, \\
\frac{\lambda \cdot \phi^\tau}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} &= \frac{\lambda_0 \cdot \phi_0^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}.
\end{aligned}
$$

Writing this out more explicitly we have, for $s = 1, ..., S$

$$
\begin{gathered}
\frac{\exp(\lambda A^s(0))}{\sum_{j \in \mathcal{S}} \exp(\lambda A^j(0))} = \frac{\exp(\lambda_0 A_0^s(0))}{\sum_{j \in \mathcal{S}} \exp(\lambda_0 A_0^j(0))} \\
\frac{\exp(\lambda A_i^s(1;\theta))}{\sum_{j \in \mathcal{S}} \exp(\lambda A_i^j(1;\theta))} = \frac{\exp(\lambda_0 A_i^s(1;\theta_0))}{\sum_{j \in \mathcal{S}} \exp(\lambda_0 A_i^j(1;\theta_0))} \\
\vdots \\
\frac{\exp(\lambda A_i^s(T;\theta))}{\sum_{j \in \mathcal{S}} \exp(\lambda A_i^j(T;\theta))} = \frac{\exp(\lambda_0 A_i^s(T;\theta_0))}{\sum_{j \in \mathcal{S}} \exp(\lambda_0 A_i^j(T;\theta_0))}.
\end{gathered}
$$

Note that for each period there are S equalities but if any $S-1$ equalities hold the remaining equality must hold because the equations sum to one. Hence we can examine $S-1$ equalities without a loss of generality. We choose the $S-1$ equalities we get after taking the ratio with respect to the first choice for each period. This result in the following set of equalities: for $s = 1, ..., S$

$$
\begin{gathered}
\frac{\exp(\lambda A^s(0))}{\exp(\lambda A^1(0))} = \frac{\exp(\lambda_0 A_0^s(0))}{\exp(\lambda_0 A_0^1(0))} \\
\frac{\exp(\lambda A_i^s(1;\theta))}{\exp(\lambda A_i^1(1;\theta))} = \frac{\exp(\lambda_0 A_i^s(1;\theta_0))}{\exp(\lambda_0 A_i^1(1;\theta_0))} \\
\vdots \\
\frac{\exp(\lambda A_i^s(T;\theta))}{\exp(\lambda A_i^1(T;\theta))} = \frac{\exp(\lambda_0 A_i^s(T;\theta_0))}{\exp(\lambda_0 A_i^1(T;\theta_0))}
\end{gathered}
$$

Define for $s = 2, ..., S$ and $t = 1, ..., T$

$$
\begin{aligned}
\Delta\pi_i(s, s_{-i}(t)) &= \pi_i(s, s_{-i}(t)) - \pi_i(1, s_{-i}(t)) \\
Z_i\left(s, s_{-i}(t)\right) &= I\{s_i(t) = s\} \cdot \pi_i(s, s_{-i}(t)) - I\{s_i(t) = 1\} \cdot \pi_i(1, s_{-i}(t)) \\
X_i\left(s, s_{-i}(t)\right) &= \Delta\pi_i(s, s_{-i}(t)) - Z_i\left(s, s_{-i}(t)\right) = I\{s_i(t) \neq s\} \cdot \Delta\pi_i(s, s_{-i}(t))
\end{aligned}
$$

Using the expression for the attraction for the initial period we have, for $s = 2, ..., S$,

$$
\lambda \left[A^s(0) - A^1(0)\right] = \lambda_0 \left[A_0^s(0) - A_0^1(0)\right].
$$

Also for the first period we have, for $s = 2, ..., S$

$$
\begin{aligned}
&\lambda \cdot \left[\frac{\phi \left[A^s(0) - A^1(0)\right] \cdot N(0) + \delta \cdot X_i(s, s_{-i}(1)) + Z_i\left(s, s_{-i}(1)\right)}{1 + \rho N(0)}\right] \\
&\qquad = \lambda_0 \cdot \left[\frac{\phi_0 \left[A_0^s(0) - A_0^1(0)\right] \cdot N_0(0) + \delta_0 \cdot X_i(s, s_{-i}(1)) + Z_i\left(s, s_{-i}(1)\right)}{1 + \rho_0 N_0(0)}\right].
\end{aligned}
$$

Analogously, in general we have, for $t = 1, ..., T$

$$\lambda \cdot \left[\frac{\phi^t \cdot \left[A^s(0) - A^1(0) \right] \cdot N(0) + \sum_{\tau=0}^{t-1} \left[\phi^\tau \cdot \delta \cdot X_i(s, s_{-i}(t-\tau)) + \phi^\tau \cdot Z_i(s, s_{-i}(t-\tau)) \right]}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} \right]$$

$$= \lambda_0 \cdot \left[\frac{\phi_0^t \cdot \left[A_0^s(0) - A_0^1(0) \right] \cdot N_0(0) + \sum_{\tau=0}^{t-1} \left[\phi_0^\tau \cdot \delta_0 \cdot X_i(s_i(t-\tau), s_{-i}(t-\tau)) + \phi_0^t \cdot Z_i(s_i(t-\tau), s_{-i}(t-\tau)) \right]}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)} \right].$$

>From the first period equality, if for each $s = 2, ..., S$

$$\left(I\{s_i(1) = s\}, X_i(s, s_{-i}(1)), Z_i(s, s_{-i}(1)) \right)$$

has a nondegenerate distribution in the area with positive probability, we have

$$\frac{\lambda \cdot \phi \left[A^s(0) - A^1(0) \right] \cdot N(0)}{1 + \rho N(0)} = \frac{\lambda_0 \cdot \phi_0 \left[A_0^s(0) - A_0^1(0) \right] \cdot N_0(0)}{1 + \rho_0 N_0(0)},$$

$$\frac{\lambda \cdot \delta}{1 + \rho N(0)} = \frac{\lambda_0 \cdot \delta_0}{1 + \rho_0 N_0(0)},$$

$$\frac{\lambda}{1 + \rho N(0)} = \frac{\lambda_0}{1 + \rho_0 N_0(0)}.$$

Analogously, if for some t

$$\left(I\{s_i(1) = s\}, X_i(s_i(1), s_{-i}(1)), Z_i(s_i(1), s_{-i}(1)), ..., I\{s_i(t) = s\}, X_i(s_i(t), s_{-i}(t)), Z_i(s_i(t), s_{-i}(t)) \right)$$

has a nondegenerate distribution in the area with positive probability, we have for $\tau = 0, ..., t-1$, $t = 1, ..., T$

$$\frac{\lambda \cdot \phi^t \left[A^s(0) - A^1(0) \right] \cdot N(0)}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} = \frac{\lambda_0 \cdot \phi_0^t \left[A_0^s(0) - A_0^1(0) \right] \cdot N_0(0)}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)},$$

$$\frac{\lambda \cdot \delta \cdot \phi^\tau}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} = \frac{\lambda_0 \cdot \delta_0 \cdot \phi^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)},$$

$$\frac{\lambda \cdot \phi^\tau}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)} = \frac{\lambda_0 \cdot \phi_0^\tau}{1 + \rho_0 + \cdots + \rho_0^{t-1} + \rho_0^t N_0(0)}.$$

It is important to recognize that the nondegeneracy of the regressors condition itself depends on the underlying behavioral model.

Since shifting the attraction times λ by a constant amount does not change the probabilities we impose the restriction that $A^1(0) = 0$. With this normalization we see that

3 Empirical implications

In this section, we discuss empirical implications of the identification results for understanding strategic behavior in games. The EWA model was developed to show that belief and reinforcement models are members of the same family of learning models. Empirically, it could provide a convenient framework that helps us to classify learning behavior in experimental games. However, as we have shown above, the framework cannot be applied without consideration of the data generating process.

We show that there are experimental data sets such that classification is not possible as none of the model's parameters is identified.

We show that there are experimental data sets such that a subset of the parameters might not be identified.

Table 1: Mookherjee and Sopher's Matching Pennies

	action R	action L
action R	$4.00, 0.00$	$0.00, 4.00$
action L	$0.00, 4.00$	$4.00, 0.00$

We show that there are experimental data sets such that all parameters are identified.

In subsection 4.1., we discuss an important class of games, games with unique mixed strategy equilibrium. When the null hypothesis of equilibrium play in Matching Pennies is not rejected, application of the EWA framework will not be informative. We show, without using the framework of the learning model, that the hypothesis of equilibrium play is not rejected in Matching Pennies experimental games (for instance Mookherjee and Sopher (1994)).

This fact has been overlooked by Camerer and Ho (1997) who had estimated the EWA learning model on Matching Pennies data. Camerer and Ho (1997) estimated the model on the data collected by Mookherjee and Sopher (1994). They state that the most striking feature of the parameter estimates is that the adjustment speed parameter λ is negative[3]. Camerer and Ho state that this might indicate the slight tendency to alternate between strategies which is a common feature in experiments with mixed strategy equilibria (p.23)[4]. Our finding is in contrast to conclusions reached in literature. Mookherjee and Sopher (1994) state that subjects are playing negatively autocorrelated strategies that are independent of one another (p.84). They provide estimated logit equations for both the population of *Row Player* subjects and the population of *Column Player* subjects. They argue that own lagged choices are significant whereas the choices of the opponents are insignificant.

In section 4.2., we review results of applications of the EWA framework to normal form games with mixed and pure equilibria and to extensive form games with multiple equilibria. In those experimental games, behavior is described well by the EWA framework. Not surprisingly, the restrictions of the belief and reinforcement learning models are generally rejected. However, it is surprising, that belief based restrictions on EWA in coordination games with pure strategy equilibria are very strongly rejected and fare less well than reinforcement learning models (contrary to Feltovich (2000)). We urge caution as towards the interpretation of the failure of the belief learning model as we may lack power to distinguish a subset of the parameters of the EWA framework when the data generating process assumes that the discount factor of experience ρ is close to zero. The subset determines the time varying speed of adjustment in EWA.

3.1 Empirical evidence on Matching Pennies

We describe Mookherjee and Sopher's Matching Pennies experiment. The game form of Matching Pennies is presented in table 1. The study was conducted with 20 fixed pairs of subjects. Each pair played the stage game repeatedly for 40 rounds. The stage game has a unique stage game equilibrium in which each player chooses each of his pure strategies with probability 0.5. Subjects were partitioned into two different groups of 10 pairs each. After each round, Group 1 subjects were given information pertaining to their own past payoffs and choices, whereas Group 2 subjects were informed in addition about the payoff matrix of the game, as well as of their opponent's past choice. Participants were paid a 20 rupees show-up fee and earned 0 or 4 rupees depending on the outcome[5]. We specialize to the data of the high information condition[6].

[3] The EWA model's estimate of the parameter is reportedly $-.43$, the reinforcement model's estimate is $-.07$ and the belief model's estimate is $-.05$ (Table 1a in Camerer and Ho (1997))

[4] Camerer and Ho also state that the estimate might indicate that subjects display sophistication: after a win, subjects expect that the opponent will switch choices and, anticipating this reaction, they themselves switch, too. We tested this conjecture, with our methodology (and hence without the EWA framework), and we found very little evidence for Camerer and Ho's hypothesis.

[5] Mookherjee and Sopher note that the average monthly expenditure of a master level student at the university at which the study was conducted was around 800 rupees.

[6] Notice that Group 1 subjects could figure out their opponent's choice in the previous round after a few games have been played and information about the possible outcomes of the game has been gathered. It is not clear what the equilibrium predictions of the

Table 2: Frequency of play, 95 % confidence interval, t statistics; Matching Pennies

Data pooled over subjects and time

mean	conf. interval	t statistic
.5125	.4778 .5472	0.7069

Data pooled over time and sorted by subject identification

i	mean	conf. interval	t statistic	i	mean	conf. interval	t statistic
1	.5000	.3381 .6619	.0000	11	.5500	.3889 .7111	.6276
2	.5500	.3889 .7111	.6276	12	.5750	.4148 .7351	.9475
3	.5000	.3381 .6619	.0000	13	.5750	.4148 .7351	.9475
4	.4250	.2649 .5851	-.9475	14	.5250	.3632 .6867	.3126
5	.5250	.3632 .6867	.3126	15	.4500	.2888 .6111	-.6276
6	.3750	.2182 .5318	-1.6125	16	.4500	.2888 .6111	-.6276
7	.7500	.6097 .8902	3.6056	17	.5000	.3381 .6619	.0000
8	.5250	.3632 .6867	.3126	18	.5500	.3889 .7111	.6276
9	.4750	.3133 .6367	-.3126	19	.5000	.3381 .6619	.0000
10	.5000	.3381 .6619	.0000	20	.4500	.2888 .6111	-.6276

We report results of tests of the conjecture that subjects follow equilibrium play in Matching Pennies. Equilibrium play implies a probability model in which choices of strategies represent independent drawings from a binomial distribution with probability .5 *and* each player's choices are independent of past choices of either player in the past[7].

Our tests are one sample t tests[8]. The sample contains 800 observations referring to a game of 10 fixed pairs of players.

3.1.1 Unconditional choice probabilities

We investigate whether choices of strategies represent drawings from a binomial distribution with the equilibrium probability. The null hypothesis is that the (unconditional) probability of choosing R equals 0.5. The alternative hypothesis is that the probability of choosing R does not equal 0.5.

Table 2 reports results of the tests. It displays the frequency of play of action R, the 95% confidence interval and the $t - statistic$ both for data pooled over time and the 20 subjects and for data pooled only over time, sorted by subject. For instance, the frequency of play of action R of the subject population is .5125, the 95% confidence interval is [.4778 .5472] and the $t - statistic$ is .7069. For instance, the frequency of play of action R of subject $i = 1$ is .5000, the 95% confidence interval is [.3381 .6619] and the $t - statistic$ is .0000.

On the data pooled over both subjects and time, the null hypothesis is not rejected in favor of the alternative hypothesis at any reasonable level of significance.

On the data pooled over time and sorted by subject identification, the hypothesis is not rejected in favor of the alternative on 18 out of 20 subjects, again at the any reasonable level of significance. The hypothesis is rejected in favor of the alternative at the 20% level of significance for a single subject and at the 1% level for the remaining subject.

---

stage and the repeated game are.

[7] We follow Brown and Rosenthal (1990) in the interpretation of the implications of equilibrium theory.

Neill (1987) reported the results of an experiment to test the minimax hypothesis for two-person, constant sum games. O'Neill interpreted the results as providing strong support for the minimax hypothesis. Brown and Rosenthal (1990) challenged this interpretation. They found that there is less evidence in support of the minimax hypothesis than indicated by O'Neill in his study and that there is strong evidence of serial correlation on player's choices.

[8] Note that the one sample t test does not require us to make distributional assumptions.

3.1.2 Conditional choice probabilities

We report results of tests of the hypothesis that subjects choose action R with probability 0.5 when we condition on choices in previous periods. The alternative hypothesis is that the probability of choosing R does not equal 0.5.

First, we investigate whether subjects display dependencies on their own past moves. Second, we investigate whether subjects display dependencies on their opponent's past moves (for the entire subject population and for the *Row Player* and *Column Player* population separately).

We condition on two kinds of variables: on subjects' own choices, I_i^s, and subjects' opponents' choices, I_{-i}^s. We condition on four kinds of histories: on previous choices, $\{I_{it-1}^s\}$ and $\{I_{-it-1}^s\}$, on the choices in two previous periods, $\{I_{it-1}^s, I_{it-2}^s\}$ and $\{I_{-it-1}^s, I_{-it-2}^s\}$, on the choices in three previous periods, $\{I_{it-1}^s, I_{it-2}^s, I_{it-3}^s\}$ and $\{I_{-it-1}^s, I_{-it-2}^s, I_{-it-3}^s\}$ and on the choices in four previous periods, $\{I_{it-1}^s, I_{it-2}^s, I_{it-3}^s, I_{it-4}^s\}$ and $\{I_{-it-1}^s, I_{-it-2}^s, I_{-it-3}^s, I_{-it-4}^s\}$.

Suppose we condition on subjects' own choices, I_i^s.

When we condition on one period lagged choices, we obtain 2 subsets of the data set The first subset contains observations if and only if the subject chose action R in the previous period i.e. $I_{it-1}^R = 1$ and the second subset contains observations if and only if the subject chose action L in the previous period i.e. $I_{it-1}^R = 0$.

When we condition on choices in two previous periods, we obtain 4 subsets of the data set. The first subset contains observations if and only if the subject chose action R in both previous periods i.e. $I_{it-1}^R = 1$ and $I_{it-2}^R = 1$, the second subset contains observations if and only if the subject chose action R in the previous period and action L two periods ago i.e. $I_{it}^R = 1$ and $I_{it-2}^R = 0$, the third subset contains observations if and only if the subject chose action L in the previous period and action R two periods ago i.e. $I_{it-1}^R = 0$ and $I_{it-2}^R = 1$ and the fourth subset contains observations if and only if the subject chose action L in both previous periods i.e. $I_{it-1}^R = 0$ and $I_{it-2}^R = 0$.

Our tests are run on those kind of subsets. Note that the data is pooled over both time and subjects.

The results are reported in tables 3 and 4. The tables display the frequency of play of action R, the 95% confidence interval and the $t - statistic$, for each subset.

Table 3 displays results from conditioning on subjects' own choices, I_i^s. For instance, when we specialize to data including observations if and only if $I_{it-1}^R = 1$, then the frequency of play of action R of the population is .4950, the 95% confidence interval is [.4456 .5443] and the $t - statistic$ is $-.2003$.

Note that

- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the history $\{I_{it-1}^s\}$
- the choice probability is significantly different from the equilibrium probability at a level of significance of 5% for 1 out of 4 tests when we condition on the histories $\{I_{it-1}^s, I_{it-2}^s\}$
- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the histories $\{I_{it-1}^s, I_{it-2,}^s I_{it-3}^s\}$
- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% level when we condition on the histories $\{I_{it-1}^s, I_{it-2}^s, I_{it-3}^s, I_{it-4}^s\}$.

We interpret these test results as evidence that subjects themselves do not reject equilibrium play as an accurate model of their own behavior.

Table 4 displays results from conditioning on subjects' opponents' choices, I_{-i}^s, for the entire population. For instance, when we specialize to data including observations if and only if $I_{-it-1}^R = 1$, then the frequency of play of action R of the population is .5251, the 95% confidence interval is [.4758 .5744] and the $t - statistic$ is 1.0025.

Table 3: Frequency of play, 95 % confidence interval, t statistic; Matching Pennies

Conditioning history: $\{I^R_{it-1}\}$

history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1\}$	.4950	.4456	.5443	-.2003	$\{0\}$	.5393	.4891	.5895	1.5377

Conditioning history: $\{I^R_{it-1},\ I^R_{it-2}\}$

history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1,1\}$	.4790	.4072	.5506	-.5794	$\{0,1\}$	.4950	.4247	.5652	-.1418
$\{1,0\}$	.5025	.4328	.5722	.0704	$\{0,0\}$	.5847	.5102	.6594	2.2437

Conditioning history: $\{I^R_{it-1},\ I^R_{it-2}, I^R_{it-3}\}$

history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1,1,1\}$	.4483	.3417	.5549	-.9645	$\{0,1,1\}$	.5361	.4351	.6371	0.7089
$\{1,1,0\}$	.5000	.3992	.6001	.0000	$\{0,1,0\}$	.4444	.3448	.5440	-1.1068
$\{1,0,1\}$	.4583	.3568	.5598	-.8151	$\{0,0,1\}$	.5980	.4986	.6973	1.9571
$\{1,0,0\}$	.5464	.4455	.6472	.9130	$\{0,0,0\}$	.5507	.4304	.6701	.8409

Conditioning history: $\{I^R_{it-1}, I^R_{it-2}, I^R_{it-3},\ I^R_{it-4}\}$

history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1,1,1,1\}$	.3611	.1963	.5259	-1.7107	$\{0,1,1,1\}$	.4681	.3199	.6161	-.4338
$\{1,1,1,0\}$	.5000	.3532	.6467	.0000	$\{0,1,1,0\}$	.5833	.4387	.7280	1.1588
$\{1,1,0,1\}$	.3902	.2343	.5461	-1.4230	$\{0,1,0,1\}$	.3921	.2534	.5308	-1.5619
$\{1,1,0,0\}$	.6037	.4677	.7399	1.5300	$\{0,1,0,0\}$	.5227	.3691	.6763	.2984
$\{1,0,1,1\}$	.5098	.3678	.6518	.1387	$\{0,0,1,1\}$	.5682	.4158	.7205	.9026
$\{1,0,1,0\}$	.3953	.2431	.5475	-1.3872	$\{0,0,1,0\}$	.6226	.4877	.7575	1.8245
$\{1,0,0,1\}$	.5345	.4021	.6667	.5219	$\{0,0,0,1\}$	.5385	.3747	.7022	.4756
$\{1,0,0,0\}$	.6000	.4292	.7707	1.1902	$\{0,0,0,0\}$	.5517	.3592	.7442	.5503

Table 4: Frequency of Play, 95 % Confidence Interval, t statistic; Matching Pennies

Conditioning history: $\{I^R_{-it-1}\}$									
history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1\}$	.5251	.4758	.5744	1.0025	$\{0\}$	.5078	.4574	.5582	.3066
Conditioning history: $\{I^R_{-it-1}, I^R_{-it-2}\}$									
history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1,1\}$	.5211	.4494	.5293	.5794	$\{0,1\}$	.5455	.4755	.6154	1.2813
$\{1,0\}$	.5274	.4577	.5970	.7751	$\{0,0\}$	.4503	.3750	.5256	-1.3027
Conditioning history: $\{I^R_{-it-1}, I^R_{-it-2}, I^R_{-it-3}\}$									
history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1,1,1\}$	.6091	.5046	.7140	2.0754	$\{0,1,1\}$	.5360	.4350	.6371	.7089
$\{1,1,0\}$	.4286	.3288	.5283	-1.4216	$\{0,1,0\}$	.5555	.4560	.6552	1.1068
$\{1,0,1\}$	.5104	.4086	.6122	.2031	$\{0,0,1\}$	.4639	.3629	.5650	-.7089
$\{1,0,0\}$	.5361	.4351	.6371	.7089	$\{0,0,0\}$	.4203	.3008	.5397	-1.3316
Conditioning history: $\{I^R_{-it-1}, I^R_{-it-2}, I^R_{-it-3}, I^R_{-it-4}\}$									
history	mean	conf. interval		t statistic	history	mean	conf. interval		t statistic
$\{1,1,1,1\}$	.6389	.4741	.8037	1.7107	$\{0,1,1,1\}$	.5319	.3838	.6800	.4338
$\{1,1,1,0\}$	.6042	.4607	.7477	1.4603	$\{0,1,1,0\}$	.5417	.3955	.6878	.5733
$\{1,1,0,1\}$	.2927	.1473	.4381	-2.8818	$\{0,1,0,1\}$	.5294	.3876	.6712	.4167
$\{1,1,0,0\}$	.5094	.3703	.6485	.1361	$\{0,1,0,0\}$	.5909	.4397	.7421	1.2125
$\{1,0,1,1\}$	.5490	.4076	.6903	.6966	$\{0,0,1,1\}$	.5455	.3923	.6986	.5986
$\{1,0,1,0\}$	.4884	.3327	.6440	-.1508	$\{0,0,1,0\}$	.3962	.2601	.5323	-1.5300
$\{1,0,0,1\}$	.5172	.3847	.6498	.2605	$\{0,0,0,1\}$	.4359	.2731	.5987	-.7969
$\{1,0,0,0\}$	.5714	.3990	.7439	.8416	$\{0,0,0,0\}$	.4138	.2231	.6044	-.9262

Table 5: Frequency of play, 95 % confidence interval, t statistic; Matching Pennies

Row Player population							
Conditioning history: $\{I^R_{it-1}\}$							
history	mean	conf. interval	t statistic	history	mean	conf. interval	t statistic
$\{1\}$	.4848	.4146 .5551	-0.4255	$\{0\}$	.5573	.4864 .6282	1.59415
Conditioning history: $\{I^R_{it-1},\ I^R_{it-2}\}$							
history	mean	conf. interval	t statistic	history	mean	conf. interval	t statistic
$\{1,1\}$	.5319	.4292 .6347	0.6168	$\{0,1\}$	.6100	.5127 .7073	2.2439
$\{1,0\}$	.4343	.3350 .5337	-1.3113	$\{0,0\}$	.4713	.3643 .5783	.5000
Conditioning history: $\{I^R_{it-1},\ I^R_{it-2}, I^R_{it-3}\}$							
history	mean	conf. interval	t statistic	history	mean	conf. interval	t statistic
$\{1,1,1\}$	.6123	.4708 .7536	1.5960	$\{0,1,1\}$	.5349	.3796 .6902	0.4532
$\{1,1,0\}$	.6123	.4708 .7536	1.5960	$\{0,1,0\}$	.5000	.3533 .6467	0.0000
$\{1,0,1\}$	.5600	.4175 .7025	0.8461	$\{0,0,1\}$	.4043	.2586 .5450	-1.3232
$\{1,0,0\}$	.3529	.1837 .5222	-1.7678	$\{0,0,0\}$	.4800	.3366 .6234	-0.2802
Column Player population							
Conditioning history: $\{I^R_{it-1}\}$							
history	mean	conf. interval	t statistic	history	mean	conf. interval	t statistic
$\{1\}$	.5650	.4957 .6343	1.8496	$\{0\}$	.4579	.3864 .5294	-1.1618
Conditioning history: $\{I^R_{it-1},\ I^R_{it-2}\}$							
history	mean	conf. interval	t statistic	history	mean	conf. interval	t statistic
$\{1,1\}$	.5104	.4086 .6122	0.2031	$\{0,1\}$	.4796	.3789 .5803	-0.4023
$\{1,0\}$	.6177	.5217 .7136	2.4330	$\{0,0\}$	.4286	.3205 .5366	-1.3150
Conditioning history: $\{I^R_{it-1},\ I^R_{it-2}, I^R_{it-3}\}$							
history	mean	conf. interval	t statistic	history	mean	conf. interval	t statistic
$\{1,1,1\}$	.6818	.5386 .8251	2.5598	$\{0,1,1\}$	.4583	.3121 .6045	-0.5733
$\{1,1,0\}$	.3600	.2222 .4978	-2.0417	$\{0,1,0\}$	.5000	.3565 .6435	.0000
$\{1,0,1\}$	.6122	.4708 .7536	1.5960	$\{0,0,1\}$	.3617	.2191 .5043	-1.9521
$\{1,0,0\}$	.5958	.4500 .7414	1.3232	$\{0,0,0\}$	.4857	.3115 .6599	-0.1667

Note that

- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the history $\{I^s_{-it-1}\}$
- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the histories $\{I^s_{-it-1}, I^s_{-it-2}\}$
- the choice probability is significantly different from the equilibrium probability at a level of significance of 5% for 1 out of 8 test when we condition on the histories $\{I^s_{-it-1}, I^s_{-it-2}, I^s_{-it-3}\}$
- the choice probability is significantly different from the equilibrium probability at a level of significance of 5% for 1 out of 16 tests when we condition on the histories $\{I^s_{-it-1}, I^s_{-it-2}, I^s_{-it-3}, I^s_{-it-4}\}$

We also test the hypothesis for the *Row Player* subjects and the *Column Player* subjects, separately. Table 5 displays results from conditioning on subjects' opponents' choices, I^s_{-i}, for the *Row Player* population in the upper panel and for the *Column Player* population in the lower panel.

Note that, for the *Row Player* population,

- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the history $\{I^s_{it-1}\}$
- the choice probability is significantly different from the equilibrium probability at a level of significance of 5% for 1 out of 4 tests when we condition on the histories $\{I^s_{it-1}, I^s_{it-2}\}$
- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the histories $\{I^s_{it-1}, I^s_{it-2,}I^s_{it-3}\}$,

  and, for the *Column Player* population,
- the choice probability is not significantly different from the equilibrium probability at a level of significance of 5% when we condition on the history $\{I^s_{it-1}\}$
- the choice probability is significantly different from the equilibrium probability at a level of significance of 5% for 1 out of 4 tests when we condition on the histories $\{I^s_{it-1}, I^s_{it-2}\}$
- the choice probability is significantly different from the equilibrium probability at a level of significance of 5% for 2 out of 8 tests when we condition on the histories $\{I^s_{it-1}, I^s_{it-2,}I^s_{it-3}\}$

We interpret these test results as evidence that subjects themselves do not reject equilibrium play as an accurate model of opponent's behavior.

3.1.3 Review of related studies

We have shown that there is strong evidence that subjects follow equilibrium play in Mookherjee and Sopher's Matching Pennies, on various level of details. We interpret the results as showing that subjects themselves do not reject equilibrium play as an accurate model of their own and of opponent's behavior.

It is important to know whether our findings generalizes to other games with unique mixed strategy equilibrium.

Elsewhere, we reanalyzed data from the careful and extensive study Roth, Erev, and Slonim (1998) undertook. Roth et al look at a universe of binary lottery games with fixed pairs of players choosing repeatedly between two actions. In those games, the equilibrium probabilities are not equal to .5. By and large, we find that there is little evidence in support of equilibrium play[9].

Bracht (2000) conducted an experimental investigation for binary lottery Matching Pennies under a fixed pairs of players matching procedure and under a random matching procedure. In our games, the equilibrium probabilities are equal to .5. (i.e. we are focusing on a smaller universe than Roth et al). We find a regularity under the fixed pairs of players matching procedure: Subjects do randomize between their choices with the equilibrium probability and do display little dependencies on past choices. Our result is in contrast to the general findings of studies of mixed equilibria on finitely iterated two person zerosum games with no pure strategy equilibria[10].

We find mixed evidence for equilibrium play under the random matching treatment.

[9] Roth et al reach the same conclusion utilizing a different methodolgy (abstracting from sample variation).

[10] For references and a statement of the general findings see Rapoport and Amaldoss (2000).

3.2 Empirical evidence on a small universe of game

The EWA learning model has been estimated on data from a wide set of games: coordination games (Camerer and Ho (1998), Camerer and Ho (1999a), Camerer and Ho (1999b)), p-beauty contests (Camerer and Ho (1999a)), games with unique mixed or pure strategy equilibrium (Camerer and Ho (1997), Camerer and Ho (1999a), Capra, Goeree, Gomez, and Holt (1999)), patent race games with iteratively dominated strategies (Rapoport and Amaldoss (2000)) and signaling games (Anderson and Camerer (2001))[11].

We think that successful applications of learning models to experimental games fall in two broad classes: data from games with unique mixed strategy equilibrium (like *Constant Sum* and *Patent Race*), data from games with multiple or unique equilibria in pure strategies (like *Weak Link*, *Median Action*, *Continental Divide* and *Traveller's Dilemma*). In addition, learning models have been successfully adapted to signaling games (*Sender-Receiver*). The model explains data in the p-beauty contests poorly[12]. Hence, we will abstract from this set.

We have argued elsewhere that maximum likelihood estimation performs well on the sample size used in practise. We concluded that to accurately estimate the parameters of the model, it is important to observe cross sectional variation. We note that there is sufficient variation in the data from the experiments we report on[13].

3.2.1 Games

We have collected all published parameter estimates of the EWA, reinforcement and belief learning models on those three classes of games. The results can be found in the appendix. Table 6 to table 8 report results from normal form games with equilibria in pure strategies. Table 9 to table 11 report results from normal form games with equilibria in mixed strategies[14]. Table 12 and table 13 report results from signaling games.

The games with pure strategy equilibria are: Median action order statistic coordination games with multiple Pareto ranked equilibria which have nine players and seven choices, the experimental game has been studied by Huyck, Battalio, and Beil (1991), estimated parameters are reported in Camerer and Ho (1997) and obtained on data with 54 subjects i.e. $I = 54$ and 10 periods i.e. $T = 10$ (*Median Action 1B*), estimated parameters are also reported in Camerer and Ho (1999b) and obtained on data with $I = 54$ and $T = 7$ (*Median Action 1A*); coordination games with two Pareto ranked equilibria which have seven players and fourteen choices, the experimental game has been studied by VanHuyck, Cook, and Battalio (1997), parameter estimates are reported in Ho, Camerer, and Chong (2001) and are obtained on data with $I = 49$ and $T = 15$ (*Continental Divide*); Traveller's Dilemma game with unique pure strategy equilibrium which have two players and multiple choices, the experimental game has been studied by Capra, Goeree, Gomez, and Holt (1999), parameter estimates are reported in Ho, Camerer, and Chong (2001) and are obtained on data with $I = 36$ and $T = 10$ (*Traveller's Dilemma*); Weak Link coordination games with multiple Pareto ranked equilibria in pure strategies which have three players and seven choices, the experimental game has been studied by Knez and Camerer (1996) and Camerer, Knez, and Weber (1996), parameter estimates are reported in Camerer and Ho (1999a) and are obtained on data with $I = 129$ and $T = 5$ (*Weak Link 2*), the experimental game has also been studied by M.Knez and Camerer (1994), parameter estimates are reported in Camerer and Ho (1997) and are obtained on data with $I = 60$ and $T = 5$ (*Weak Link 1*).

The games with unique mixed strategy equilibria are: Constant sum games with mixed equilibrium studied by Mookherjee and Sopher (1997) which have four (*Constant Sum 1* and *2*) or six choices (*Constant Sum 3* and *4*), one of which is weakly dominated, parameter estimates are reported in Camerer and Ho (1997) and are obtained on data with $I = 20$ and $T = 28$, for each game; Patent Race games with unique mixed equilibrium

[11] There are some extensions of the EWA model. For example, Camerer, Hsia, and Ho (2000) modifies EWA to account for behavior in bilateral call markets.

[12] See ?.Subsequently, in ?, Camerer et al extend the EWA model. They report that the extended EWA improves the fit considerably.

[13] See Ichimura and Bracht (2001) for an examination of the performance of statistical techniques used to estimate learning models on experimental game data.

[14] * parameter estimates are not significantly different from zero (at a level .01)
** parameter estimates are significantly different from zero (at a level of .01)

which have two players either with the same endowment and six choices (*Patent Race 1 and 2*) or with a different endowment and six choices for the strong player and five choices for the weak player (*Patent Race 3*), parameter estimates are reported in Rapoport and Amaldoss (2000) and are obtained on data with $I = 36$ and $T = 80$, for each game.

The signaling games are: Sender-Receiver games (1 and 2) with multiple equilibria in pure strategies which have a sender of two types t_1 and t_2, two messages m_1 and m_2 and a receiver with three actions a_1, a_2 and a_3. Attractions for senders are type dependent and attractions for receivers are message dependent. For the sender, attractions to messages are updated for the unrealized type. For the receiver, attractions to actions are updated for the unchosen actions. Parameter estimates are reported in Anderson and Camerer (2001) and are obtained on data with $I = 36$ and $T = 24$.

3.2.2 Restriction on parameter space

We wish to review applications of the EWA framework. Recall that for identification, one of the initial attractions needs to be normalized. In the earlier studies i.e. in estimation of the model on data from *Median Action 1B*, *Weak Link 1* and *2*, the parameter space was not restricted. In later studies, in estimation of the model on data from *Median Action 1A*, *Constant Sum 1-4* and *Patent Race 1-3*, the parameter space was restricted as follows: $\phi > 0$, $\lambda > 0$, $\delta \geq 0$, $\rho \leq 1$, $A_i^s(0)$ is smaller than difference between maximum of the possible payoffs minus the minimum of the possible payoffs and $0 \leq N(0) \leq (1-\rho)^{-1}$. In estimation of the model on data from *Continental Divide* and the *Traveller's Dilemma*, the initial attractions correspond to choice probabilities that match the actual population frequency of choices in the first period. The parameter space was restricted as follows: $N(0) \leq (1-\rho)^{-1}$. In estimation on data from signaling games, the parameter space was restricted as follows: $A_i^s(0)$ is larger than the minimum possible payoff and smaller than the maximum possible payoff, $0 \leq N(0) \leq (1-\rho)^{-1}$, $N(0) < 50$ and $0 < \rho < 1$.

3.2.3 A taxonomy of estimation results

In this subsubsection, we aim to classify learning behavior in experimental games, using the EWA framework[15]. We do so by consulting the parameter estimates across games and by deciding whether behavior is more of a belief type or more of a reinforcement type.

Next, we will review how successful the restricted and unrestricted EWA models explain the data, by consulting the fit to the data. Not surprisingly, the unrestricted model generally does better than the restricted cases. We are interested in the question in which sense is actual behavior different from behavior described and predicted by the restricted cases.

Recall that the reinforcement learning restrictions on the general model are $\delta = 0$, $\rho = 0$ and $N(0) = 1$ and that the belief learning restrictions are $\delta = 0$ and $\rho = \phi$. Recall also the the important imagination parameter δ is almost always identified. This result allows us to classify behavior.

First, we have a look at estimates on data from games with mixed equilibrium. For the EWA model, there is some dispersion of the estimates of δ. The estimates are low and are found in the interval [.00 .73], the estimates of ϕ are close to 1 and found in the interval [.90 1.04] and the estimates of ρ are found in the interval [.92 .98]. Note that the estimates of the two discount parameters are close. For the belief model, the estimate of ρ ($=\phi$) are very close to 1 and found in the interval [.99 1.00]. For the reinforcement model, the estimates of ϕ are close to 1, too, and found in the interval [.93 1.01].

Second, we have a look at estimates on data from three games with pure equilibrium. For the EWA model, there is some dispersion of the estimates of δ. The estimates are high and are found in the interval [.53 .853], the estimates of ϕ are found in the interval [.74 .80] and the estimates of ρ are 0. Note that the estimates of the two discount parameters are not close. For the belief model, the estimate of ρ ($=\phi$) are closer to 1 and found

[15] Salmon (2001) simulation analysis indicates that successful classification depends somewhat on the precise specification of the data generating process. Hence, we might obtain less accurate results than we obtained in Ichimura and Bracht (2001).

in the interval [.738 .99]. For the reinforcement model, the estimate of ϕ is found at .93[16]. We also have a look at estimates on data from three other games with pure equilibrium. Note that the estimates reported in the left and right panels are somewhat different. In the panel to the right, the estimates of δ and ϕ tend to be lower and the estimates of ρ tend to be larger. We think this is due to different set of restrictions imposed in the studies.

Third, we have a look at estimates on data from extensive form games with incomplete information (signaling games). For the EWA model, the estimates of δ are .69 and .54, the estimates of ϕ are 1.02 and .65, and the estimates of ρ are 1.00 and .46. Note that the estimates of the two discount factors are not close. For the belief learning model the estimates of $\rho (= \phi)$. are .97 and .88[17].

We tentatively conclude that there is a mapping from experimental games to estimated parameters of the framework. In mixed normal form games, the estimates of the imagination parameter δ tend to be low and the estimates of the discount parameters, ρ and ϕ, tend to be close to 1. In pure normal form games, the estimates of the imagination parameter δ tend to be high and the estimates of the discount parameter ρ are (close to) 0. In extensive form games with incomplete information, estimates are at a medium level for δ[18].

3.2.4 Statistical fit and interpretation of the failure of the special cases

In the previous subsubsection, we have provided a mapping from parameter estimates to games. As the estimates of the framework are not all equal to the restricted parameters of either special case, we conjecture that the EWA framework does a better job characterizing behavior in the selected experimental games. Note that the EWA framework is all but one of many hybridizations of the special cases. Hence, it will prove its validity if and only if its application is informative about the reasons for the failure of the special cases.

We report the statistical fit of the 3 models. In table 14, for each model, for the normal form games, we report the negative likelihood and a statistic testing the restrictions of the special cases. In table 16, for the normal form games, we report the negative likelihood for the predictions of the models on a holdout sample. Note that the restrictions imposed by the special cases are generally rejected[19]. For the signaling games, we report the average per period likelihood (summed across subjects) for the first 24 periods of the experiment in table ??. This is the number that was minimized in estimation. The second and fourth row present the same statistics for the 25th to 32th periods, the holdout sample into which we hope to predict.

First, we look at the statistical fit in games with unique mixed strategy equilibrium. The belief model explains the data from *Constant Sum* almost as well as the EWA model. This is surprising as the estimates of the imagination parameter δ are low (except for game 2). The estimates of the initial count of observations, $N(0)$, in EWA are high and close to the steady state values. Hence, there is not much change in the frequency of choices of the strategies, for each of the four constant sum games. The belief model is able to capture this feature by estimating a very high initial count of observations, $N(0)$, for each of the four constant sum games. This interpretation is confirmed by the results from the holdout sample (the belief model does better than EWA in 2 out of 4 cases). In contrast, the belief model does a bad job on the data from *Patent Race*. This is not surprising as the estimates of the imagination parameter δ are low.

The reinforcement model does a bit worse than the EWA model. On data from *Constant Sum*, it does worse than the belief model. The reinforcement model captures the slow change in the choice frequencies by estimating the speed adjustment parameter λ close to 0. The reinforcement learning model does much better than the belief

[16] For the Traveller's Dilemma and the Continental Divide, Camerer et al investigate a variant on the reinforcement model.

[17] Note that the discount factor of experience is restricted to be less than one such that ρ and $N(0)$ are not identified, on data from game 1. Note that data from game 2, the restriction Camerer and Ho impose is violated by the reported estimates.

Estimation results on the reinforcement learning model are not available.

[18] Please note that we have collected *all* the published estimates of the framework. Clearly, further research is needed to enlarge the set of applications.

[19] An exception is the constant sum game 2. The belief model is not rejected in favor of the EWA model.

learning model on the *Patent Race* data and almost as well as the EWA model. Note that it appears that subjects do not utilize cumulatives yet averages of past and prior payoff experience (as the framework's estimates of ρ are far from zero).

We now turn to estimation results on data from pure games. The EWA model does a good job capturing and predicting behavior in those games. The reinforcement model tends to fail to predict switching to actions that are best responses (recall that the framework's estimates of the imagination parameter are close to 1 and so the reinforcement models failure is hardly surprising). The belief model tends to fail to predict the sharp convergence of play.

Hence, in sample and out of sample, the special cases are clearly rejected in favor of the general model (see 14 and 16).

The failure of the belief learning model is surprising. So, a few words are in order to explain what is happening. We specialize to the Median action data set. Feltovich (2000) noted that belief learning is substantially better than reinforcement learning on the median action data set.

We will urge caution towards the interpretation of the failure of the belief learning model, in *Median Action* and *Weak Link*.

To illustrate, we follow the description in Camerer and Ho (1999b).

Consider the *Median Action*. In this game, players pick a number between 1 and 7 and they earn a payoff that decreases in the deviation from the median of all players' action. An equilibrium is not observed in the first period of a sequence of play. The initial choices are concentrated around $4-5$, with a dip at 6 and small spikes at 3 and 7. At the end of the sequence, the equilibrium was always equal to the median of the distribution of actions chosen in the first period of that sequence. Choices of 3 and 7 are quickly extinguished. The EWA model fits the patterns of choices well. It captures the frequencies of initial choices, the increase in choices of $4-5$ and best responding. The reinforcement learning model fails to explain the switching away from 7 fast enough. The belief learning model does not place enough weight on 7, as 7 is chosen more frequently, and does not capture the steady increase in $4-5$ at all.

Reinforcement learning does not capture the increase in $4-5$ fast enough, as it does not capture best responding (as δ is restricted to be 0). As a result, initial attractions are a composite between the first period data and later period data.

Belief learning does capture best responding (as δ is restricted to be 1 and the EWA estimate of δ is close to 1), however it fails.

We argue that the model might fail for a similar reason reinforcement learning fails. The belief model fails as it does not capture initial choices since initial attractions are restricted to be expected payoffs based on prior beliefs. As a result, initial attractions are a composite between the first period data and later period data.

In contrast, Camerer and Ho argue that the combination of parameter estimates $\delta < 1$ and $\phi > \rho \simeq 0$ is crucial for the success of the EWA model as this combination assures sharp convergence. They argue that the belief learning model fails to capture convergence as it restricts the imagination parameter $\delta = 1$ and, with the restriction $\phi = \rho$, ρ does not equal 0.

We urge caution towards Camerer and Ho's interpretation. Note that the estimate of the discount factor of experience $\rho = 0$ for median action is zero. Hence, it might well be that actual play follows reinforcement learning s.t. $\rho = 0$. If that would be the case, the parameters ρ, $N(0)$ and λ are not separately identified. In addition, the initial attractions, $A_i^s(0)$, are not separately identified, either. Consequently, we cannot tell whether the interpretation Camerer and Ho assign to the estimates is valid or not. It might very well be that belief model fails as initial play is restricted, whereas in unrestricted EWA (and reinforcement learning) initial play is not restricted.

Finally, we look at the statistical fit in signaling games. The belief learning model does worse than the EWA framework. This is not surprising as the estimate of the imagination parameter δ is not close to 1. As a consequence, behavior described and predicted by belief learning is too sluggish.

4 Conclusion

The paper makes basic yet important contributions.

First, we provide identification results of the experience weighted attraction learning model and its special cases, reinforcement and belief learning.

Second, we use the EWA framework to identify learning behavior in experimental games. Our identification results allow us to do so, under consideration of the data generating process. We discuss two empirical implications of the identification results.

We show that, on data from games with unique mixed strategy equilibrium, the EWA specification does not allow identification of learning models when players follow equilibrium play. We show that subjects indeed follow equilibrium play in experimental Matching Pennies, contrary the conclusions reached in the literature.

Nevertheless, we argue that the EWA framework allows us to classify behavior in almost all experimental games as the most important parameter distinguishing members of the model's family is identified, under some regularity conditions. We tentatively provide a classification of learning behavior in games. Behavior in games with unique mixed strategy equilibrium is described well by reinforcement type models in which little weight is given to foregone payoffs. Behavior in games with multiple equilibria might be described well by belief type models in which much weight is given to foregone payoffs.

This is not the first study to identify learning behavior in experimental games.

Erev and Roth (1998) analyzed a large set of data from games with unique mixed strategy equilibrium. The authors compare a number of learning models by fitting the models to the data and testing which models fits best. They found that reinforcement learning does a good job on the data. While we use a different methodology, and classify behavior according to the estimates of the general model, we do conform their main result.

Feltovich (2000) analyzed data from games with multiple equilibria. Feltovich compares a number of learning models by fitting the models to the data and testing which models fits best, in a fashion similar to Erev and Roth (1998). Feltovich looks at models that are all special cases of the general model. Feltovich found that belief learning does a better job than reinforcement learning fitting the data from coordination games, contrary to results of goodness of fit presented in this paper. Note that Feltovich (2000) attempted to limit the reliance of the models on initial attractions by *not* comparing predictions of behavior with actual play in early rounds. However, in the studies reviewed in this paper, restriction on initial expected payoffs in belief learning are imposed. For this reason, we think, belief learning does much worse than reinforcement learning, in the studies on coordination games reviewed. Belief learning predicts sluggish behavior during the course of play as the model is forced to put a structure on first period play. EWA and reinforcement learning are not forced to do so, and both predict better.

Camerer and Ho made a significant contribution by showing that, while belief and reinforcement learning models are conceptually different, behaviorally, the models are commensurable in the sense that a general learning model encompasses the two kinds of models as special cases. It is an empirical question whether the EWA model, in its current specification, is a valid hybrid. We have shown that there are identification problems of the general model. Identification problems in the reinforcement and belief models are not always alleviated by introducing the parameters ρ, $N(0)$ and δ in EWA (falsifying claims inCamerer and Ho (1999b) on *p*.868). On the contrary, for Matching Pennies, when play follows equilibrium, none of the parameters of the model is identified and, for coordination games, when play is modeled well by reinforcement learning, a subset of the parameters is not identified.

5 Appendix

Proof of Theorem 1: Recall that the updating rule of observation equivalent is

$$N(t) = 1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)$$

and that the updating rules of attractions for $t \geq 1$ is specified is

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

Defining $X_i^s(t) = A_i^s(t) \cdot N(t)$, we have

$$X_i^s(t) = \phi \cdot X_i^s(t-1) + [\delta + (1-\delta) \cdot I\{s_i(t) = s\}] \cdot \pi_i(s, s_{-i}(t)),$$

so that for $t \geq 1$

$$X_i^s(t) = \phi^t \cdot X_i^s(0) + \sum_{\tau=0}^{t-1} \phi^\tau \cdot [\delta + (1-\delta) \cdot I\{s_i(t-\tau) = s\}] \cdot \pi_i(s, s_{-i}(t-\tau))$$

and that for $t \geq 1$

$$A_i^s(t) = \frac{\phi^t \cdot A^s(0) \cdot N(0) + \sum_{\tau=0}^{t-1} \phi^\tau \cdot [\delta + (1-\delta) \cdot I\{s_i(t-\tau) = s\}] \cdot \pi_i(s, s_{-i}(t-\tau))}{1 + \rho + \cdots + \rho^{t-1} + \rho^t N(0)}.$$

References

Anderson, C., and C. Camerer (2001): "Experience Weighted Attraction Learning in Sender-Receiver Signalling games," *Economic Theory*, 16, 689–718.

Bracht, J. (2000): "Empirical Investigation of Learning Models," Ph.D. thesis, University of Pittsburgh, Pittsburgh, U.S.A.

Brown, J. N., and R. W. Rosenthal (1990): "Testing the Minimax Hypothesis: A re-examiniation of O'Neills Game Experiment," *Econometrica*, 58, 1065–1081.

Camerer, C., and T. Ho (1997): "Experience-Weighted Attraction Learning in Games: A unifying approach," Caltech, Social Science Working Paper, 1003.

Camerer, C., and T. Ho (1998): "EWA Learning in Games: Heterogeneity and Time Variation," *Journal of Mathematical Psychology*, 42, 305–326.

——— (1999a): "Experience-Weighted Attraction Learning in Games: Estimates from Weak-Link Games," in *Games and Human Behavior, Essays in Honor of Amnon Rapoport*, ed. by D. Budescu, I. Erev, and R. Zwick. Kluwer Academic, Dordrecht/Norwell, M.A.

——— (1999b): "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica*, 67, 827–874.

Camerer, C., D. Hsia, and T. Ho (2000): "EWA Learning in Bilateral Call Markets," .

Camerer, C., M. Knez, and R. Weber (1996): "Timing and virtual observability in ultimatum bargaining and weak-link coordination games," Caltech, Social Science Working Paper, 970.

Capra, M., J. Goeree, R. Gomez, and C. Holt (1999): "Anomalous Behavior in a Traveler's Dilemma," *American Economic Review*, 89.

Erev, I., and A. Roth (1998): "Predicting How People Play Games: Reinforcement Learning in Experimental Games With Unique, Mixed Strategy Equilibria," *American Economic Review*, 88, 848–881.

Feltovich, N. (2000): "Reinforcement-Based vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games," *Econometrica*, 68, 605–642.

Ho, T., C. Camerer, and J. Chong (2001): "Economic Value of EWA Lite: A Functional Theory of Learning in Games," Caltech, Social Science Working Paper 1122.

Huyck, J. V., R. Battalio, and R. Beil (1991): "Strategic Uncertainty, Equilibrium Selection and Coordination Failure in Averae Opinion Games," *Quarterly Journal of Economics*, 106, 885–909.

Ichimura, H., and J. Bracht (2001): "Estimation of Learning Models on Experimental Game Data," Center for Rationality, Working Paper No.243.

Knez, M., and C. Camerer (1996): "Increasing cooperation in social dilemmas through the precedent of efficiency in coordination games," University of Chicago Working Paper.

M.Knez, and C. Camerer (1994): "Creating expectational assets in the laboratory: Coordination in weakest-link games," *Strategic Management Journal*, 15, 101–119.

Mookherjee, D., and B. Sopher (1994): "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 7, 62–91.

——— (1997): "Learning and Decision Costs in Experimental Constant Sum Games," *Games and Economic Behavior*, 19, 97–132.

Neill, B. (1987): "Nonmetric test of the minimax theory of two person zero sum games," *Proceedings of the National Academy of Sciences U.S.A.*, 84, 2106–2109.

Rapoport, A., and W. Amaldoss (2000): "Mixed Strategies and Iterative Elimination of Strongly Dominated Strategies: An experimental investigation of states of knowledge," *Journal of Economic Behavior and Organization*, 1, 1–1.

Roth, A. E., I. Erev, and R. Slonim (1998): "Learning and Equilibrium as Useful Approximations: Accracy of Prediction on Randomly Selected Constant Sum Games," Mimeo, University of Pittsburgh.

Salmon, T. (2001): "An Evaluation of Econometric Models of Adaptive Learning," *Econometrica*, 69, 1597–1629.

VanHuyck, J., J. Cook, and R. Battalio (1997): "Adaptive Behavior and Coordination Failure," *Journal of Economic Behavior and Organization*, 32, 483–503.

Appendix: Published estimation results of the EWA, reinforcement and belief learning model on experimental game data

Table 6: Parameter estimates of the EWA learning model on data from games with unique pure strategy equilibrium and games with multiple equilibria (Standard errors in parentheses, jacknife procedure)

Parameters	Median Action 1A	Continental Divide	Traveller's Dilemma	Weak Link 1	Weak Link 2	Median Action 1B
δ	.853	.73	.53	.652	.66	.90
	(.005)	(.09)	(.07)	*n.a.*	*n.a.*	*n.a.*
ϕ	.800	.74	.77	.582	.53	.54
	(.018)	(.03)	(.02)	*n.a.*	*n.a.*	*n.a.*
ρ	.000	.00	.00	.198	.22	.55
	(.000)	(.02)	(.02)	*n.a.*	*n.a.*	*n.a.*
λ	6.827	3.98	3.53)	6.157	5.94	33.49
	(.251)	(.30)	(.22)	*n.a.*	*n.a.*	*n.a.*
$N(0)$	.647	.25	.62	2.187	3.78	9.64
	(.059)	(.00)	(.00)	*n.a.*	*n.a.*	*n.a.*
$A^1(0)$	.000	–	–	.763	.69	.90
	–			*n.a.*	*n.a.*	*n.a.*
$A^2(0)$	.000	–	–	.656	.59	.91
	(.145)			*n.a.*	*n.a.*	*n.a.*
$A^3(0)$	.337	–	–	.765	.72	1.01
	(.144)			*n.a.*	*n.a.*	*n.a.*
$A^4(0)$	.561	–	–	.834	.77	1.01
	(.143)			*n.a.*	*n.a.*	*n.a.*
$A^5(0)$	.505	–	–	.716	.75	1.00
	(.143)			*n.a.*	*n.a.*	–
$A^6(0)$	.364	–	–	.658	.75	.98
	(.148)			*n.a.*	*n.a.*	*n.a.*
$A^7(0)$	.431	–	–	1.00	1.00	.99
	(.146)			–	–	*n.a.*

Table 7: Parameter estimates of the belief learning model on data from games with unique pure strategy equilibrium and games with multiple equilibria (Standard errors in parentheses, jacknife procedure)

Parameters	Median Action 1A	Continental Divide	Traveller's Dilemma	Weak Link 1	Weak Link 2	Median Action 1B
$\phi = \rho$	.738	.99	.85	.525	.36	.36
	(.054)	(.05)	(.01)	*n.a.*	*n.a.*	*n.a.*
λ	16.726	14.76	13.97	21.98	13.15	24.60
	(.366)	(1.26)	(.85)	*n.a.*	*n.a.*	*n.a.*
$N(0)$	.642	1.11	6.74	11.08	16.71	4.38
	(.150)	(.15)	(1.51)	*n.a.*	*n.a.*	*n.a.*
$A^1(0)[N^1(0)]$	.168[.075]	–	–	.700[4.71]	.70[5.32]	.33
	(.007)(.018)			*n.a.*	*n.a.*	*n.a.*
$A^2(0)[N^2(0)]$	.491[.049]	–	–	.715[.00]	.74[.00]	.64
	(.005)(.012)			*n.a.*	*n.a.*	*n.a.*
$A^3(0)[N^3(0)]$	.713[.059]	–	–	.730[.00]	.77[.05]	.85
	(.002)(.014)			*n.a.*	*n.a.*	*n.a.*
$A^4(0)[N^4(0)]$	.836[.073]	–	–	.745[.00]	.81[.00]	.97
	(.003)(.019)			*n.a.*	*n.a.*	*n.a.*
$A^5(0)[N^5(0)]$	.859[.093]	–	–	.760[.00]	.84[.00]	.98
	(.003)(.022)			*n.a.*	*n.a.*	*n.a.*
$A^6(0)[N^6(0)]$	.781[.123]	–	–	.755[.00]	.88[.00]	.90
	(.006)(.028)			*n.a.*	*n.a.*	*n.a.*
$A^7(0)[N^7(0)]$	.604[.170]	–	–	.791[6.38]	.92[11.33]	.79
	(.008)(.038)			*n.a.*	*n.a.*	*n.a.*

Table 8: Parameter estimates of the reinforcement learning model on data from games with unique pure strategy equilibrium and games with multiple equilibria (Standard errors in parentheses, jacknife procedure)

Parameters	Median Action 1A	Weak Link 1	Weak Link 2	Median Action 1B
ϕ	.930	.527	.50	.87
	(.009)	*n.a.*	*n.a.*	*n.a.*
λ	1.190	2.319	2.00	1.02
	(.015)	*n.a.*	*n.a.*	*n.a.*
$A^1(0)$	.000	1.153	.09	−4.08
	–	*n.a.*	*n.a.*	*n.a.*
$A^2(0)$	.000	.806	−.36	−3.37
	(.028)	*n.a.*	*n.a.*	*n.a.*
$A^3(0)$	.877	1.119	.25	−.834
	(.033)	*n.a.*	*n.a.*	*n.a.*
$A^4(0)$	2.400	1.253	.45	1.18
	(.000)	*n.a.*	*n.a.*	*n.a.*
$A^5(0)$	2.261	.746	.21	1.00
	(.021)	*n.a.*	*n.a.*	–
$A^6(0)$	1.329	.358	−.03	−.41
	(.033)	*n.a.*	*n.a.*	*n.a.*
$A^7(0)$	.986	1.565	1.00	−.78
	(.041)	*n.a.*	–	*n.a.*

Table 9: Parameter estimates of the EWA learning model on data from seven games with unique mixed strategy equilibrium (standard errors in parentheses, jacknife procedure)

	Constant Sum				Patent Race			
Parameters	1	2	3	4	1	2	3	
δ	.000	.730	.413	.547	.000	.000	.475*	.137*
	(.035)	(.103)	(.082)	(.054)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
ϕ	1.040	1.005	(.986)	.991	.940*	.970*	.901*	.959*
	(.010)	(.009)	(.005)	(.011)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
ρ	.961	.946	.935	.926	.929*	.984*	.857*	.973*
	(.014)	(.011)	(.006)	(.024)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
λ	.508	.182	.646	.218	.7555*	.967*	.4911*	1.644*
	(.048)	(.015)	(.030)	(.019)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$N(0)$	19.630	18.391	15.276	9.937	4.613*	15.747*	1.418*	12.454*
	(.065)	(.713)	(.009)	(.017)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
Row Player					Both Players		Strong	Weak
$A^1(0)$	1.320	5.237	2.996	9.491	.000	.000	.000	1.724*
	(.059)	(.222)	(.068)	(.145)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^2(0)$	1.145	1.172	2.434	7.964	.000	.000	7.696*	.000
	(.092)	(.265)	(.043)	(.090)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^3(0)$	1.913	7.187	.000	.027	.000	.000	7.432*	.491 **
	(.066)	(.300)	(.015)	(.081)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^4(0)$	.000	.000	.000	.004	.000	.000	7.120*	.951 **
	–	–	(.036)	(.001)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^5(0)$	–	–	1.338	6.105	.000	.000	9.188*	1.534*
	–	–	(.034)	(.060)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^6(0)$	–	–	.000	.000	1.929*	2.301*	11.681*	–
	–	–	–	–	*n.a.*	*n.a.*	*n.a.*	–
Column Player								
$A^1(0)$	2.681	6.790	4.998	8.733	–	–	–	–
	(.091)	(.640)	(.020)	(.198)	–	–	–	–
$A^2(0)$	2.583	6.859	4.047	6.306	–	–	–	–
	(.102)	(.580)	(.046)	(.080)	–	–	–	–
$A^3(0)$	2.345	5.359	2.201	1.572	–	–	–	–
	(.092)	(.735)	(.050)	(.049)	–	–	–	–
$A^4(0)$	.000	.000	3.852	6.565	–	–	–	–
	–	–	(.043)	(.070)	–	–	–	–
$A^5(0)$	–	–	1.737	1.851	–	–	–	–
	–	–	(.047)	(.267)	–	–	–	–
$A^6(0)$	–	–	.000	.000	–	–	–	–
	–	–	–	–	–	–	–	–

Table 10: Parameter estimates of the belief learning model on data from seven games with unique mixed strategy equilibrium (standard errors in parentheses, jacknife procedure)

Parameters	Constant Sum				Patent Race			
	1	2	3	4	1	2	3	
$\phi = \rho$	1.000	1.000	.989	1.000	1.000	1.000	.999∗	1.000
	(.0001)	(.0005)	(.004)	(.002)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
λ	1.168	.459	1.812	1.501	.475∗	.149∗	1.440∗	3.726∗
	(.067)	(.063)	(.123)	(.019)	*n.a*	*n.a.*	*n.a.*	*n.a.*
$N(0)$	300.000	58.000	90.000	30.000	303.295∗	233.758∗	162.382∗	312.914∗
	(.000)	(.001)	(.012)	(.448)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
Row Player					Both Players		Strong	Weak
$A^1(0)[N^1(0)]$	1.780[107.370]	3.914[22.703]	2.309[41.566]	5.335[16.005]	.000	.000	.000	.76∗
	(.029)(1.758)	(.073)(.4230	(.045)(.810)	(.053)(.283)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^2(0)[N^2(0)]$	1.422[107.290]	1.739[25.214]	2.077[11.048]	4.665[.0001]	.000	.000	1.668∗	.000
	(.070)(4.687)	(.290)(1.403)	(.036)(.886)	(.053)(.000)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^3(0)[N^3(0)]$	2.262[85.335]	4.927[10.084]	1.122[10.227]	.734[1.090]	.000	.000	.0812∗	.152
	(.057)(4.187)	(0.149)(1.682)	(.086)(1.111)	(.050)(.083)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^4(0)[N^4(0)]$	.948[.000]	1.159[.000]	1.182[18.018]	.364[10.704]	.092∗	.000	1.412∗	.411
	(.047)(.000)	(.193)(.000)	(.080)(.850)	(.026)(.298)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^5(0)[N^5(0)]$	–	–	1.615[9.141]	3.568[2.200]	.000	16.359∗	1.541∗	.555∗
	–	–	(.030)(.993)	(.071)(.115)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^6(0)[N^6(0)]$	–	–	1.076[.000]	.097(.000)	3.116∗	16.359∗	2.545∗	–
	–	–	(.038)(.000)	(.040)(.000)	*n.a.*	*n.a.*	*n.a.*	–
Column Player								
$A^1(0)[N^1(0)]$	3.385[96.887]	6.674[19.294]	3.595[25.296]	7.769[6.692]	–		–	
	(.049)(2.923)	(.092)(.536)	(.044)(.797)	(.072)(.253)	–		–	
$A^2(0)[N^2(0)]$	3.079[87.932]	6.423[17.962]	3.218[32.627]	6.383[10.301]	–		–	
	(.043)(2.582)	(.089)(.935)	(.034)(1.151)	(.131)(.496)	–		–	
$A^3(0)[N^3(0)]$	2.896[115.280]	5.711[20.744]	2.444[16.539]	3.954[5.141]	–		–	
	(.040)(2.563)	(.136)(.514)	(.068)(.922)	(.102)(.352)	–		–	
$A^4(0)[N^4(0)]$	1.281[.001]	2.384[.000]	3.068[13.391	6.557[5.684	–		–	
	(.029)(.000)	(.059)(.000)	(.053)(.914)	(.150)(.462)	–		–	
$A^5(0)[N^5(0)]$	–	–	2.269[2.147]	4.135[.028]	–		–	
	–	–	(.085)(.389)	(.213)(.000)	–		–	
$A^6(0)[N^6(0)]$	–	–	1.663[.000]	3.608[2.155]	–		–	
	–	–	(.034)(.000)	(.131)(.373)	–		–	

Table 11: Parameter estimates of the reinforcement learning model on data from seven games with unique mixed strategy equilibrium (standard errors in parentheses, jacknife procedure)

Parameters	Constant Sum				Patent Race			
	1	2	3	4	1	2	3	
ϕ	1.012	.978	.960	.962	.927∗	.954∗	.925∗	.941∗
	(.006)	(.008)	(.005)	(.005)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
λ	.053	.033	.098	.046	.066∗	.030∗	.0345∗	.063∗
	(.004)	(.002)	(.005)	(.002)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
Row Player					Both Players		Strong	Weak
$A^1(0)$	.000	10.000	5.000	10.000	.000	.000	.000	47.789∗
	(.012)	(.000)	(.000)	(.000)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^2(0)$	.000	.000	5.000	10.000	.000	.000	164.925∗	.000
	(.000)	(.000)	(.000)	(.000)	*n.a.*	*n.a..*	*n.a.*	*n.a.*
$A^3(0)$	5.000	10.000	.000	.000	.000	.000	146.647∗	13.433∗
	(.000)	(.000)	(.000)	(.000)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^4(0)$	.000	.000	.000	.000	.000	.000	147.987∗	26.167∗
	–	–	(.000)	(.000)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^5(0)$	–	–	.229	8.501	.000	.000	166.002∗	39.225∗
	–	–	(.000)	(.000)	*n.a.*	*n.a.*	*n.a.*	*n.a.*
$A^6(0)$	–	–	.000	.000	19.810∗	74.635∗	193.4721∗	–
	–	–	–	–	*n.a.*	*n.a.*	*n.a.*	–
Column Player								
$A^1(0)$	5.000	10.000	5.000	10.000				
	(.000)	(.000)	(.000)	(.000)				
$A^2(0)$	5.000	10.000	4.852	9.996				
	(.000)	(.000)	(.000)	(.000)				
$A^3(0)$	5.000	8.408	.001	.001				
	(.000)	(.000)	(.000)	(.000)				
$A^4(0)$	.000	.000	4.973	9.832				
	–	–	(.000)	(.000)				
$A^5(0)$	–	–	.000	.001				
	–	–	(.000)	(.000)				
$A^6(0)$	–	–	.000	.000				
	–	–	–	–				

Table 12: Parameter estimates of the EWA learning model on data from two signaling games (bootstrapped 95% confidence interval in parentheses)

	Sender-Receiver	
Parameters	1	2
δ	.69	.54
	$(.47, 1.00)$	$(.45, .63)$
ϕ	1.02	.65
	$(.99, 1.04)$	$(.59, .71)$
ρ	1.00	.46
	$(.98, 1.04)$	$(.39, .54)$
λ	.41	.09
	$(.34, .54)$	$(.07, .11)$
$N_S^{m_1}(0) = N_R^{m_1}(0)$	50.00	.62
	$(49.91, 50.00)$	$(.59, .66)$
$N_S^{m_2}(0) = N_R^{m_2}(0)$	32.91	3.37
	$(32.81, 32.94)$	$(3.37, 3.38)$
$A_S^{m_1 t_1}(0)$	15.00	18.25
	–	$(18.25, 18.26)$
$A_S^{m_2 t_1}(0)$	9.90	30.00
	$(9.04, 10.72)$	–
$A_S^{m_1 t_2}(0)$	14.72	30.00
	$(13.94, 15.24)$	–
$A_S^{m_2 t_2}(0)$	15.00	11.34
	–	$(11.34, 11.34)$
$A_R^{a_1 m_1}(0)$	30.00	30.00
	–	–
$A_R^{a_2 m_1}(0)$	25.12	37.26
	$(24.61, 25.74)$	$(37.26, 37.26)$
$A_R^{a_3 m_1}(0)$	15.00	0.00
	$(15.00, 15.00)$	$(.00, .01)$
$A_R^{a_1 m_2}(0)$	20.08	41.88
	$(19.37, 20.32)$	$(41.88, 41.88)$
$A_R^{a_2 m_2}(0)$	21.78	15.00
	$(21.16, 21.87)$	–
$A_R^{a_3 m_2}(0)$	15.00	43.26
	–	$(43.26, 43.26)$

Table 13: Parameter estimates of the belief learning model on data from two signaling games (95% confidence interval in parentheses)

	Sender-Receiver						
Parameters	1			2			
$\phi = \rho$	.97				.88		
	(.97, .99)				(.88, .94)		
λ	.32				.15		
	(.27, .38)				(.12, .18)		
$N_S^{m_1}(0) = N_R^{m_1}(0)$	34.64	$N_S^{a_1m_1}(0)$	.00	$N_S^{m_1}(0) = N_R^{m_1}(0)$	2.84	$N_s^{a_1m_1}(0)$	2.18
			(.00, .00)				(2.17, 2.20)
$N_S^{m_2}(0) = N_R^{m_2}(0)$	26.07	$N_S^{a_2m_1}(0)$	10.40	$N_S^{m_2}(0) = N_R^{m_2}(0)$	8.12	$N_s^{a_2m_1}(0)$	.00
			(10.39, 10.41)				(.00, .01)
$A_S^{m_1t_1}(0)$	25.50	$N_S^{a_3m_1}(0)$	24.24	$A_S^{m_1t_1}(0)$	34.70	$N_s^{a_3m_1}(0)$	.65
			(24.23, 24.24)				(.62, .67)
$A_S^{m_2t_1}(0)$	17.56	$N_S^{a_1m_2}(0)$	.00	$A_S^{m_2t_1}(0)$	44.62	$N_s^{a_1m_2}(0)$	3.60
			(.00, .00)				(3.58, 3.61)
$A_S^{m_1t_2}(0)$	20.99	$N_S^{a_2m_2}(0)$	15.90	$A_S^{m_1t_2}(0)$	30.00	$N_s^{a_2m_2}(0)$	.56
			(15.89, 15.90)				(.54, .59)
$A_S^{m_2t_2}(0)$	20.85	$N_S^{a_3m_2}(0)$	10.17	$A_S^{m_2t_2}(0)$	19.95	$N_s^{a_3m_2}(0)$	3.96
			(10.17, 10.18)				(3.95, 3.97)
$A_R^{a_1m_1}(0)$	30.00	$N_R^{m_1t_1}(0)$	17.39	$A_R^{a_1m_1}(0)$	30.00	$N_R^{m_1t_1}(0)$	.63
			(17.39, 17.39)				(.58, .65)
$A_R^{a_2m_1}(0)$	22.40	$N_R^{m_2t_1}(0)$	6.88	$A_R^{a_2m_1}(0)$	35.04	$N_R^{m_2t_1}(0)$	3.06
			(6.86, 6.88)				(3.04, 3.07)
$A_R^{a_3m_1}(0)$	15.00	$N_R^{m_1t_2}(0)$	17.25	$A_R^{a_3m_1}(0)$	3.32	$N_R^{m_1t_2}(0)$	2.21
			(17.24, 17.25)				(2.20, 2.23)
$A_R^{a_1m_2}(0)$	23.76	$N_R^{m_2t_2}(0)$	19.19	$A_R^{a_1m_2}(0)$	33.93	$N_R^{m_2t_2}(0)$	5.06
			(19.19, 19.20)				(5.05, 5.08)
$A_R^{a_2m_2}(0)$	26.04			$A_R^{a_2m_2}(0)$	24.35		
$A_R^{a_3m_2}(0)$	15.00			$A_R^{a_3m_2}(0)$	31.96		

Table 14: In sample fit of three models of learning on data from 14 games

Experimental Game	EWA $-LL$	Belief $-LL$	Belief χ^2	Reinforcement $-LL$	Reinforcement χ^2
Constant Sum 1	653.072	680.232	54.320	681.968	57.792
Constant Sum 2	678.496	681.632	5.328	710.136	63.280
Constant Sum 3	790.608	797.720	14.224	853.608	126.000
Constant Sum 4	855.288	842.968	21.280	901.600	92.264
Patent Race 1	3551.70	4649.39	2195.386	3563.76	24.111
Patent Race 2	2908.08	3634.02	1451.890	2928.29	40.416
Patent Race 3	3031.54	3738.31	1413.542	3120.60	178.116
Patent Race 4	2835.51	3446.49	1221.953	2851.16	31.310
Median Action 1999b	309.304	438.748	258.88	258.88	64.785
Continental Divide	1045	1279.1	–	–	–
Traveller's Dilemma	856.287	1106.2	–	–	–
Median Action 1997	344.03	522.15	356.24	406.80	125.54
Weak Link 1997	358.058	438.546	160.98	386.787	58.22
Weak Link 1999a	813.99	983.24	338.50	852.37	76.94

Table 15: Out of sample fit of three models of learning on data from 7 games

Experimental Game	EWA $-LL$	Belief $-LL$	Reinforcement $-LL$
Constant Sum 1	326.38	285.60	335.50
Constant Sum 3	341.71	350.09	359.74
Constant Sum 2	301.70	296.28	308.47
Constant Sum 4	362.26	371.18	375.94
Median Action 1999b	41.05	113.90	80.27
Continental Divide	460	565	–
Traveller's Dilemma	443	465	–

Table 16: In and out of sample fit of two models of learning on data from 2 signaling games

Signaling Game	EWA $-LL$	Belief $-LL$
In sample Game 1	12.06	12.55
Out of sample Game 1	9.02	10.28
In sample Game 2	13.49	16.41
Out of sample Game 2	17.44	20.400