

A comparison of learning and replicator dynamics using experimental data

Yin-Wong Cheung, Daniel Friedman*

Economics board, University of California, Santa Cruz, CA 95064, USA

Received 16 September 1997; received in revised form 21 October 1997; accepted 21 October 1997

Abstract

We compare the explanatory power of two popular adjustment models, replicator dynamics and simple belief learning dynamics. The data, gathered over 18 laboratory sessions each with 8 to 16 profit-motivated players, comprise more than 100 runs (each with 10 or more stages or periods) of single and two population 2×2 bimatrix games with unique interior equilibria. To equalize the number of free parameters we suppress individual player and treatment effects known to be highly significant in the simple belief learning model. Nevertheless the belief learning model narrowly outperforms the replicator model on the single population data and greatly outperforms it on the two population data. © 1998 Elsevier Science B.V.

JEL classification: C72; C73; C92

Keywords: Belief learning; Experimental data; Mixed Nash equilibrium; Replicator dynamics

1. Introduction

Game theorists recently have used adjustment dynamics to explain how Nash equilibrium is achieved and to explain selection among multiple Nash equilibria; see (Weibull, 1995; Fudenberg and Levine, 1998) for recent textbook presentations. Many different kinds of adjustment models are theoretically plausible but different models can produce quite different predictions, so empirical work is in order (Crawford, 1995b). In this paper we use laboratory data to compare empirically two of the most popular models in the theoretical literature, replicator dynamics and simple belief learning.

* Corresponding author.

Replicator dynamics (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1988) arises naturally from population games in biology. It postulates that the ‘market share’ or relative prevalence of any strategy has a growth rate proportional to its payoff relative to the average payoff. These dynamics approximate certain specifications of rote learning (Börgers and Sarin, 1993) and imitation (Schlag, 1994).

Belief learning models postulate that players adjust their beliefs as they accumulate experience, and that current beliefs influence the current choice of strategy. Cournot (1838), for example, assumes that players noiselessly best respond to the belief that other players’ current actions will be the same as in the previous period. Fictitious play (Brown, 1951) assumes that players noiselessly best respond to the belief that other players’ current actions will be equal to the average of their actions in all earlier periods. Crawford (1995a), Cheung and Friedman (1997), denoted CF97 below, consider simple belief learning models that reduce to Cournot or fictitious play for particular parameter values.

In this paper we offer empirical specifications for replicator dynamics (as far as we know, for the first time) and for a simple learning model with the same number of free parameters. We fit the models to laboratory data from binary choice normal form games with a unique, interior equilibrium. Some of the games are completely symmetric (single population) while other games have two distinct player populations or types, analogous to row players and column players in bimatrix games. Using comparisons that give replicator dynamics their ‘best shot,’ we find that the learning model is slightly better at explaining the single population data and much better at explaining the two population data.¹

The rest of the paper is organized as follows. Section 2 introduces notation and contains a brief discussion of population games in normal form. Section 3 presents the replicator dynamics model and the three parameter belief learning model. A summary description of the laboratory experiments and data appears in Section 4. Section 5 reports the empirical results, and Section 6 concludes.

2. Population games

We study how the distribution of behavior in player populations evolves over time when players interact via simple 2×2 bimatrix games. The formal structure is a variant of a repeated normal form game with N players. At each stage $t=1, 2, \dots, T$ of the game, each player $i=1, \dots, N$ chooses an action a_{it} from the set A_i . The payoff to an individual player is represented by a function $G_i : A_i \times A_{-i} \rightarrow R$. That is, the stage t payoff to player i in general depends on her own action a_{it} and the profile $a_{-it} \in A_{-i} \equiv A_1 \times \dots \times A_{i-1} \times A_{i+1} \times \dots \times A_N$ of other players’ actions.

In this paper we consider only binary choice games in which each player’s action set $A_i=\{1,0\}$ has the same two actions, and each player faces a single strategically homogeneous population of n opponents. By strategic homogeneity, we mean that the

¹ A few preliminary results appear in the Spanish language article Friedman (1993). The experiment is presented more fully in Friedman (1996) and the learning model is presented and tested more fully in Cheung and Friedman (1997). The underlying set of experiments include games with multiple or pure Nash Equilibria. We restrict our attention here to the cleanest test case, games with a unique mixed Nash equilibrium.

state variable is the opponents’ action distribution. With binary choice, the state variable s_{it} for player i at time t is simply fraction of current opponents choosing the first action; i.e. $s_{it} = n^{-1} \sum_{j=1}^n a_{jt}$, where a_{jt} is the action chosen by opponent j .

Two kinds of population games are considered. In a *single population* (or *symmetric*) game, each player has $n=N-1$ opponents and the payoff function $G_i(a_{it}, a_{-it}) = g(a_{it}, s_{it})$ is the same for every player $i=1, \dots, N$. For example, the symmetric two-player bimatrix $\begin{pmatrix} -2, -2 & 8, 0 \\ 0, 8 & 4, 4 \end{pmatrix}$ can be reinterpreted as a single population stage game in which the payoff to $a_{it}=1$ (i.e., ‘row’ plays ‘top’) is $g(1, s_{it}) = -2s_{it} + 8(1-s_{it}) = 8 - 10s_{it}$ and the payoff to $a_{it}=0$ (‘row’ plays ‘bottom’) is $g(0, s_{it}) = 4 - 4s_{it}$. Note that in a single population game s_{it} differs from the aggregate state $s_t = N^{-1} \sum_{j=1}^N a_{jt}$ only by excluding own action from the average.

In a *two population game* (referred to in the biology literature as ‘the asymmetric’ case), the players are divided into two groups ($k=1, 2$) of equal size $n=N/2$. Players in one group are matched only against players in the other, so here the relevant state s_{it} is the fraction of players in the other group choosing the first action. The payoff function in this case is $G_i(a_{it}, a_{-it}) = g_k(a_{it}, s_{it})$, which may depend on which group k that player i belongs to but otherwise is independent of i . For example, the two-player bimatrix $\begin{pmatrix} 2, 2 & 0, -1 \\ 3, 3 & 1, 4 \end{pmatrix}$ can be reinterpreted as a two population stage game in which payoff for the first (‘row’) population of $a_{it}=1$ (‘top’) is $g_1(1, s_{it}) = 2s_{it}$, and for the second (‘column’) population the payoff to $a_{it}=1$ (‘left’) is $g_2(1, s_{it}) = 2s_{it} + 3(1-s_{it})$, etc.

We work with payoff functions g (or g_k) that are linear in the sense that $g(a, s) = sg(a, 1) + (1-s)g(a, 0)$. In the single population case, the matrix representation is $M = (m_{ij})_{i,j=1,2}$ where $m_{ij} = g(2-i, 2-j)$, since the second action is coded as 0 rather than 2. For the two population case, the payoff matrix for group $k=1, 2$ is the 2×2 matrix $M_k = ((m_{ijk}))_{i,j=1,2}$ with components $m_{ijk} = g_k(2-i, 2-j)$.² One can verify by direct calculation that for a player’s mixed strategy $(P, 1-P)$, the expected payoff in a single population game is $(P, 1-P)M(s_{it}, 1-s_{it})'$, where P is the player’s probability to choose $a_{it}=1$, and s_{it} is the state variable. If a two population game is considered, the expected payoff becomes $(P, 1-P)M_k(s_{it}, 1-s_{it})'$ where now s_{it} is the average action in the other population.

Linear binary choice single population games come in three basic types distinguished according to how the advantage of the first action, say, depends on the state s . Define the advantage

$$\begin{aligned} R(s) &= s(g(1, 1) - g(0, 1)) + (1 - s)(g(1, 0) - g(0, 0)) \\ &= (1, -1)M(s, 1 - s)' \\ &= b + cs \end{aligned} \tag{1}$$

where $b \equiv (1, -1)M(0, 1)'$ and $c \equiv (1, -1)M(1, -1)' = m_{11} + m_{22} - m_{21} - m_{12} \equiv g(1, 1) + g(0, 0) - g(0, 1) - g(1, 0)$. Then an interior steady state of the single population game is defined as a solution $s^* \in (0, 1)$ to $R(s) = 0$.

² Note that our notation differs slightly from the usual textbook 2×2 bimatrix notation $(M_1; M_2')$, where group 1 is the ‘row’ players and group 2 is the ‘column’ players, and the prime denotes transpose. In particular, we do not transpose the group 2 matrix M_2 .

The game is called type 1 if $R(s)=0$ has a solution $s^*=-b/c$ in $(0, 1)$ and the slope c of R is negative; type 2 if s^* is in $(0, 1)$ and c is positive; and type 3 if s^* lies outside $(0, 1)$. We ignore the case in which $b=c=0$ since then payoffs are constant, independent of own and others' actions, and the game is trivial; the knife-edge cases $s^*=0$ or 1 are also neglected here for simplicity. Friedman (1996), denoted F96 below, shows that there is a dominant action and hence a unique corner steady state (and a corresponding unique pure Nash equilibrium of the associated symmetric bimatrix game) in type 3 games; an interior steady state at s^* and two corner steady states at $s=0$ and $s=1$ (with three corresponding NE of the associated symmetric bimatrix game) in type 2 games; and a unique steady state at s^* (and a corresponding mixed NE of the associated symmetric bimatrix game) in type 1 games.³

It is not difficult to show that replicator and learning models (and most other adjustment dynamics) make the same prediction for behavior in type 2 and type 3 single population games: direct convergence to the corner steady state corresponding to the dominant strategy NE in type 3 games, and direct convergence to the corner steady state on the same side of s^* as the current state in type 2 games. Type 1 is the crucial case for distinguishing between the models, and we will focus our theoretical and empirical work on this case in single population games and in their two population analogues.

In a two population game, the advantage $R_k(s)$ for group $k=1, 2$ is defined using Eq. (1) with M_k , b_k and c_k replacing M , b , and c . Recall that here a player's payoff depends on behavior in the other population, so an interior steady state is a solution pair $(s_1^*, s_2^*) \in (0, 1) \times (0, 1)$ to the equations $R_k(s)=0$, $k=2, 1$. The mixture $(s_1^*, 1-s_1^*)$ by the row player and $(s_2^*, 1-s_2^*)$ by the column player is a mixed Nash equilibrium of the associated bimatrix game precisely when $R_1(s_2^*)=0=R_2(s_1^*)$.

3. Population dynamics

The relevance of Nash equilibrium of the bimatrix game depends on whether behavior in the population game actually converges to a steady state; i.e. it depends on the adjustment dynamics. In this section we present the two most popular models of adjustment dynamics and derive some theoretical properties.

3.1. Replicator dynamics

Replicator dynamics postulates that the action distribution adjusts over time with 'market share' growth rates for each action proportional to its expected payoff relative to the population average. For a single population game with 2×2 matrix M , replicator dynamics reduces to

$$\begin{aligned} \Delta s_{t+1}/s_t &= \beta \{ (1, 0)M(s_t, 1-s_t)' - (s_t, 1-s_t)M(s_t, 1-s_t)' \} \\ &= \beta(1-s_t)R(s_t) \end{aligned} \quad (2)$$

³ In the last case, type 1, the associated symmetric bimatrix game also has two asymmetric pure NE, but asymmetric NE have no corresponding steady states in single population games. As explained in F96 p. 7 and Weibull (1995), these asymmetric NE do have corresponding steady states in two population games but these two population games are type 2 analogues, not type 1 analogues.

where $\Delta s_{t+1} = s_{t+1} - s_t$ is the change in the state variable, $(1 - s_t)R(s_t)$ is the relative payoff of action 1, and $\beta > 0$ is the adjustment speed. A few words of interpretation and literature comparison are in order. Eq. (2) derives the relative payoff variable $(1 - s)R(s)$ as the payoff to the first action minus the population-weighted average payoff, with weights s and $(1 - s)$ on the two actions. By contrast, the (absolute) advantage variable $R(s)$ is the payoff to the first action minus the payoff to the second action.

The relative payoff is used directly in the original continuous time version of replicator dynamics (Taylor and Jonker, 1978). Discrete time and other versions often normalize by dividing the relative payoff by the expression $k + (s, 1 - s)M(s, 1 - s)$ for some free parameter $k \geq 0$ (e.g. Binmore et al., 1995; Weibull, 1995). The normalization is needed in some discrete time versions to ensure that market shares sum to 1, but is unnecessary in our binary choice case. Some authors include the normalization even when unnecessary (for example, to reflect an unmodelled background payoff as noted in Vega-Redondo, 1996) but most (including us) do not, for the sake of theoretical simplicity.⁴

When will single population replicator dynamics produce convergence to an interior equilibrium s^* ? Write (2) as $\Delta s_{t+1} = \beta s_t(1 - s_t)R(s_t)$ and Taylor expand both sides of the equation around s^* . Recall from the previous section that $R(s^*) = 0$ and $R'(s) = c$. Up to first order in $\tilde{s}_t = s_t - s^*$, we have

$$\begin{aligned} \Delta \tilde{s}_{t+1} &= 0 + \tilde{s}_t \beta [\partial \{s_t(1 - s_t)R(s_t)\} / \partial s_t |_{s_t=s^*}] \\ &= \tilde{s}_t \beta c s^*(1 - s^*) \end{aligned}$$

Hence, the replicator dynamics locally is governed by the linearization

$$\tilde{s}_{t+1} = \phi \tilde{s}_t$$

where $\phi = 1 + \beta c s^*(1 - s^*)$. We have local convergence iff $|\phi| < 1$, and oscillations around s^* iff $|\phi| < 0$.⁵ Hence, when $0 < c \equiv (1, -1)M(1, -1)'$, we have $\phi > 1$ and the interior Nash equilibrium s^* is replicator-unstable, but when $0 > c$, we have direct convergence if $0 < \beta \leq \beta_s \equiv [c | s^*(1 - s^*)]^{-1}$, oscillatory convergence if $\beta_s < \beta < 2\beta_s$, and nonconvergence if $\beta > 2\beta_s$. Replicator dynamics in the two population game expresses the growth rate in population $k (= 1, 2)$ as

$$\begin{aligned} \Delta s_{kt+1} / s_{kt} &= \beta_k \{ (1, 0)M_k(s_{jt}, 1 - s_{jt})' - (s_{kt}, 1 - s_{kt})M_k(s_{jt}, 1 - s_{jt})' \} \\ &= \beta_k (1 - s_{kt})R(s_{jt}) \end{aligned} \tag{3}$$

where $(1, 0)M_k(s_{jt}, 1 - s_{jt})'$ is the expected payoff to the first action, $(s_{kt}, 1 - s_{kt})M_k(s_{jt}, 1 - s_{jt})'$ is the population k average expected payoff, and $j = 3 - k$ is the opponent population. That is,

$$\Delta s_{1t+1} = \beta_1 s_{1t} (1 - s_{1t}) R_1(s_{2t})$$

⁴ We also have an empirical reason for omitting the normalization. Early replicator regressions on our data convincingly showed that unnormalized versions fit the data better than normalized versions. The normalized fits became comparable to the unnormalized fits only when k was larger than 4. But for k this large, the variation in $(s, 1 - s)M(s, 1 - s)$ has little effect and the normalization is essentially only a rescaling of the replicator variable. For this reason, only estimates based on Eq. (2), the unnormalized replicator dynamics, are presented in Section 4.

⁵ Here we ignore the borderline case $|\phi| = 1$ where higher order terms determine stability because this case has probability zero in our setting. See Hirsch and Smale (1974) for a good textbook treatment.

and

$$\Delta s_{2t+1} = \beta_2 s_{2t} (1 - s_{2t}) R_2(s_{1t}) \quad (4)$$

It turns out that replicator dynamics is never stable at an interior Nash equilibrium (s_1^*, s_2^*) . If $c_1 = R_1'(s)$ and $c_2 = R_2'(s)$ have opposite signs then (s_1^*, s_2^*) is a neutral center, i.e. the state (s_1, s_2) circles around (s_1^*, s_2^*) , neither converging nor diverging, and it is a saddle point if $c_1 c_2 > 0$.⁶

3.2. A belief learning model

Our other empirical model arises from a theoretical account of players' decision process and belief revision process presented more fully in CF97. Suppose player i 's expectation \hat{s}_{it+1} of the opponents' action distribution in the next period is based on her observations so far of the actual states s_{it-u} ; $u=0, \dots, t-1$. If she regards recent observations as more informative (and rightly so since the state is adjusting) and uses a discount factor γ , then we have

$$\hat{s}_{it+1} = \left[s_{it} + \sum_{u=1}^{t-1} \gamma_i^u s_{it-u} \right] / \left[1 + \sum_{u=1}^{t-1} \gamma_i^u \right] \quad (5)$$

Eq. (5) encompasses two classical models of learning from experience. If $\gamma=0$, it reduces to the Cournot expectations (Cournot, 1838) with $\hat{s}_{it+1} = s_{it}$. If $\gamma=1$, we have fictitious play (Brown, 1951) where all the previous state observations have the same weight, $\hat{s}_{it+1} = [\sum_{u=1}^t s_{t-u}] / t$. If $0 < \gamma < 1$, we have adaptive expectations model in which more recent observations have a greater weight. Values of $\gamma > 1$ imply that older observations have a greater weight as in 'imprinting'. Values of $\gamma < 0$ are highly counterintuitive in that they imply that the influence of a given observation changes sign each period.

For given \hat{s}_{it} , the expected advantage to the first action is $R(\hat{s}_{it})$ from Eq. (2). The proposed decision rule is stochastic with the probability that individual i selects action 1 given by

$$P(a_{it} = 1) = F(\alpha_i + \beta_i R(\hat{s}_{it})) \quad (6)$$

where $F(\cdot)$ is the standard normal cumulative distribution function,⁷ β_i captures the player's responsiveness to the perceived advantage, and α_i is a catch-all intercept parameter that, among other things, reflects the player's idiosyncratic preference for the first action.

If all players obey this learning-decision model (and have identical parameters) in a single population game, then the actual state s_{t+1} will have approximately the binomial

⁶ The basic idea can be seen by calculating J , the Jacobian of (4), and noting that its main diagonal terms are 0 and the off-diagonal terms are $\beta_k c_k s_k^* (1 - s_k^*) \equiv d_k$, with the same sign as c_k , for $k=1, 2$. Hence the eigenvalues of the linearized difference equation $(\tilde{s}_{1t+1}, \tilde{s}_{2t+1})' = (J + I)(\tilde{s}_{1t}, \tilde{s}_{2t})'$, where $\tilde{s}_{it} = s_{it} - s_i^*$, are $\lambda = 1 \pm (d_1 d_2)^{1/2}$. See Hofbauer and Sigmund (1988) and (Weibull (1995), 5.2.2) for more sophisticated treatments.

⁷ In early stages of the exercise, we found no essential differences in using the logit distribution function instead of the normal. We use normal distribution mainly for its familiarity and convenience in interpreting coefficients.

distribution (N, P) where $P = F(\alpha + \beta R(\hat{s}_t))$. The approximation ignores the difference between s_{it-u} and the aggregate state s_{t-u} . Up to first order near an interior Nash equilibrium s^* , we have

$$Es_{t+1} = F(\alpha + \beta R(s^*)) + (s_t - s^*)[\partial\{F(\alpha + \beta R(\hat{s}_t))\}/\partial\hat{s}_t |_{\hat{s}_t=s^*}] \quad (7)$$

Stability of s^* (in the sense that if historical states s_{iu} ; $u=1, \dots, t$, are all ϵ -close to s^* then Es_{t+1} is even closer) hinges on both α and β . If $\alpha = \alpha^* \equiv F^{-1}(s^*)$, then the first term in (7) is $F(\alpha^* + \beta R(s^*)) = F(\alpha^*) = s^*$. In this case, there is no tendency to leave s^* once there, and Eq. (7) reduces to

$$\tilde{s}_{t+1} = \theta \tilde{s}_t$$

where now \tilde{s} is the expected deviation from s^* and

$$\begin{aligned} \theta &= \partial\{F(\alpha^* + \beta R(\hat{s}_t))\}/\partial\hat{s}_t |_{\hat{s}_t=s^*} \\ &= F'(\alpha^*)\beta c \end{aligned}$$

We have stability if $|\theta| < 1$, i.e. $|\beta| < (F'(\alpha^*)|c|)^{-1} \leq (2\pi)^{1/2}/|c|$. The last inequality follows from the fact that $F'(\cdot)$ is the standard normal density function, whose maximum is $(2\pi)^{-1/2}$. See the Appendix of CF97 for a variant of this stability argument.

For two population games we again need $\alpha_k = \alpha_k^* = F^{-1}(s_k^*)$; $k = 1, 2$, to keep the state from systematically drifting away from an interior Nash equilibrium (s_1^*, s_2^*) . Linearizing the system at the Nash equilibrium gives

$$Es_{kt+1} = F(\alpha_k + \beta_k R(s_j^*))$$

for $k=1, 2$ and $j=3-k$. This leads to a transition matrix for the expected deviation \tilde{s}_{kt} that has zeros on the main diagonal and terms $d_k = F'(\alpha_k^*)\beta_k c_k$ off the diagonal. The eigenvalues are $\pm (d_1 d_2)^{1/2}$. The system is stable if their modulus is less than one; i.e. if $|d_1 d_2| < 1$ or $|\beta_1 \beta_2| < [F'(\alpha_1^*)F'(\alpha_2^*)|c_1 c_2|]^{-1} \leq 2\pi/|c_1 c_2|$.

4. Data

The data are drawn from the laboratory experiment described in F96. The single population data consist of observations from experiments using the Hawk–Dove

matrix $M = \begin{pmatrix} -2 & 8 \\ 0 & 4 \end{pmatrix}$. The two population data are generated from the bimatrix

$\left[\begin{pmatrix} 2 & 0 \\ 3 & -1 \end{pmatrix}; \begin{pmatrix} 2 & 3 \\ -1 & 4 \end{pmatrix} \right]$ in which population 1 is called the Buyers and population 2

the Sellers. Population sizes range from $N=8$ to $N=16$. Most runs in these experiments repeat the stage game for $T=10$ periods but some repeat for up to $T=16$ periods. Players typically complete 6 or more runs and earn \$ 5–\$ 20 in each session.

A review of the key treatments will help in interpreting the data. The experiments employ two alternative procedures for matching players. Under the *random pairwise* (RP) procedure, the computer randomly picks a matching scheme independently in each period, each admissible scheme being equally likely. For a 2×2 matrix M , a player's

expected payoff is AMS' if he chooses pure strategy $A=(a, 1-a)$ and the action distribution by potential opponents is $S=(s, 1-s)$. However, his actual payoff depends on the action taken by his actual opponent, and so has some variance around its expectation. The variance is eliminated in the alternative matching procedure, called *mean matching* (MM). Here each player is matched once against each possible opponent in each round and receives the average (mean) payoff, either $(1, 0)MS'$ for action $a=1$ or $(0,1)MS'$ for action $a=0$.

A: Random Pairwise Matching

- Average # of (A,B,C): (7.2,4.8,0) -					
round	Total (A,B,C)	Your Choice	Payoff		
1	(9,3,0)	A	-2.0	Y	
2	(8,4,0)	A	-2.0	O	
3	(8,4,0)	A	-2.0	U	
4	(4,8,0)	B	4.0		
5	(7,5,0)	B	0.0		
Your total earnings: -2.00					
Please choose 'a' or 'b':					

Other Player	
A	B
A	B
-2.0	8.0
B	B
0.0	4.0

Each period the other player you face is randomly chosen from the other 11 players in your group.

B: Mean Matching

- Average # of (A,B,C): (5.6,6.4,0) -				# of A's	A payoff	B payoff
round	Total (A,B,C)	Your Choice	Payoff			
1	(5,7,0)	A	3.8	0	8.0	4.0
2	(4,8,0)	B	2.7	1	7.2	3.7
3	(8,4,0)	A	1.3	2	6.3	3.3
4	(6,6,0)	A	3.0	3	5.5	3.0
5	(5,7,0)	B	2.3	4	4.7	2.7
				5	3.8	2.3
				6	3.0	2.0
				7	2.2	1.7
				8	1.3	1.3
				9	0.5	1.0
				10	-0.3	0.7
				11	-1.2	0.3
				12	-2.0	0.0
Your total earnings: 13.1						
Please choose 'a' or 'b':						

Fig. 1. Players' screens.

Another treatment variable is the amount of historical information to appear on each player's computer screen. In the minimum level, *No History*, the player receives no historical information other than what she could tabulate herself: her own action and actual payoff in previous periods. In the usual level, *History*, the screen lists the actual states of the relevant population in previous periods.

Fig. 1 shows players' computer screens in experiments with *History* treatment. RP matching is displayed in Panel A and MM is in Panel B. In either case, the player enters and confirms his action; 'a' for the first action ($a_{it}=1$) and 'b' for the second action ($a_{it}=0$). The payoff matrix M appears on the right side of the screen. The chosen row is highlighted. When all players have confirmed their actions, the realized payoff is highlighted in the payoff box and is also displayed on the upper left box along with other historical data. In the case of *No History*, the historical data box is eliminated from the screen. Interested readers are referred to F96 for a more detailed discussion of the procedures.

5. Empirical results

Before fitting the competing models, we offer a graphical introduction to the data, borrowed from F96. Fig. 2 shows the first four Hawk–Dove runs. Some readers may detect a tendency for the state to return to the interior NE (shown as a horizontal line) via stable oscillations and perhaps less noise with MM (runs 3 and 4) than with RP (runs 1 and 2). Other readers will remain skeptical until they see more data and systematic analysis. Similarly, Fig. 3 shows the first four Buyer–Seller runs; less skeptical readers may detect neutrally stable cycles especially in the first two MM runs (1 and 2).

5.1. The learning model

CF97 estimates the learning model separately for each player using the specific historical states s_{it-u} and allows the coefficients to vary with the history and matching treatments. Using likelihood ratio tests, they show that fits are very significantly poorer in specifications that impose identical coefficients restrictions across individuals or across treatments. Nevertheless, we will impose the restrictions here in order to keep the number of free parameters the same as in the replicator model. For the single population (Hawk–Dove) data, we set $\alpha_i=\alpha$, $\beta_i=\beta$, and $\gamma_i=\gamma$ for every player i , $i=1, \dots, N$, and we use the aggregate state variable $s_t = N^{-1} \sum_{j=1}^N a_{jt}$. Thus, for $t>1$ the parameters are estimated using the equations

$$\begin{aligned} P(a_{it} = 1) &= F(\alpha + \beta R(\hat{s}_t)) \\ R(\hat{s}_t) &= (1, -1)M(\hat{s}_t, 1 - \hat{s}_t) \\ \hat{s}_{t+1} &= \left[s_t + \sum_{u=1}^{t-1} \gamma^u s_{t-u} \right] / \left[1 + \sum_{u=1}^{t-1} \gamma^u \right] \end{aligned} \quad (8)$$

where α is the 'average' tendency to select the first action and β is the 'average' response to expected payoff differential, which is determined by the past mean state variables and the 'average' memory parameter γ . Recall that $F(\cdot)$ is the standard normal distribution

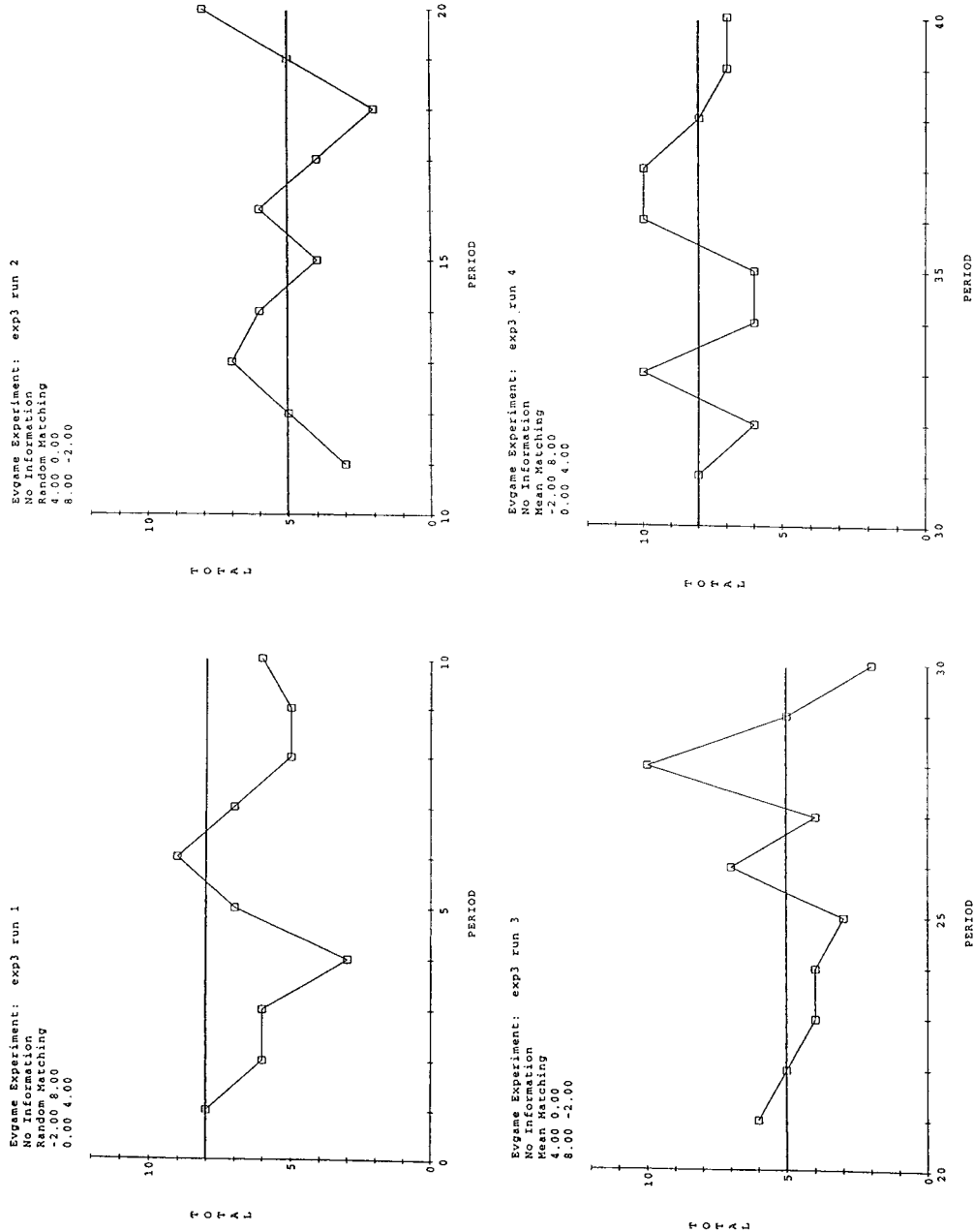


Fig. 2. Session Exp 3, runs 1–4. Note: The equilibrium is indicated by the horizontal bar.

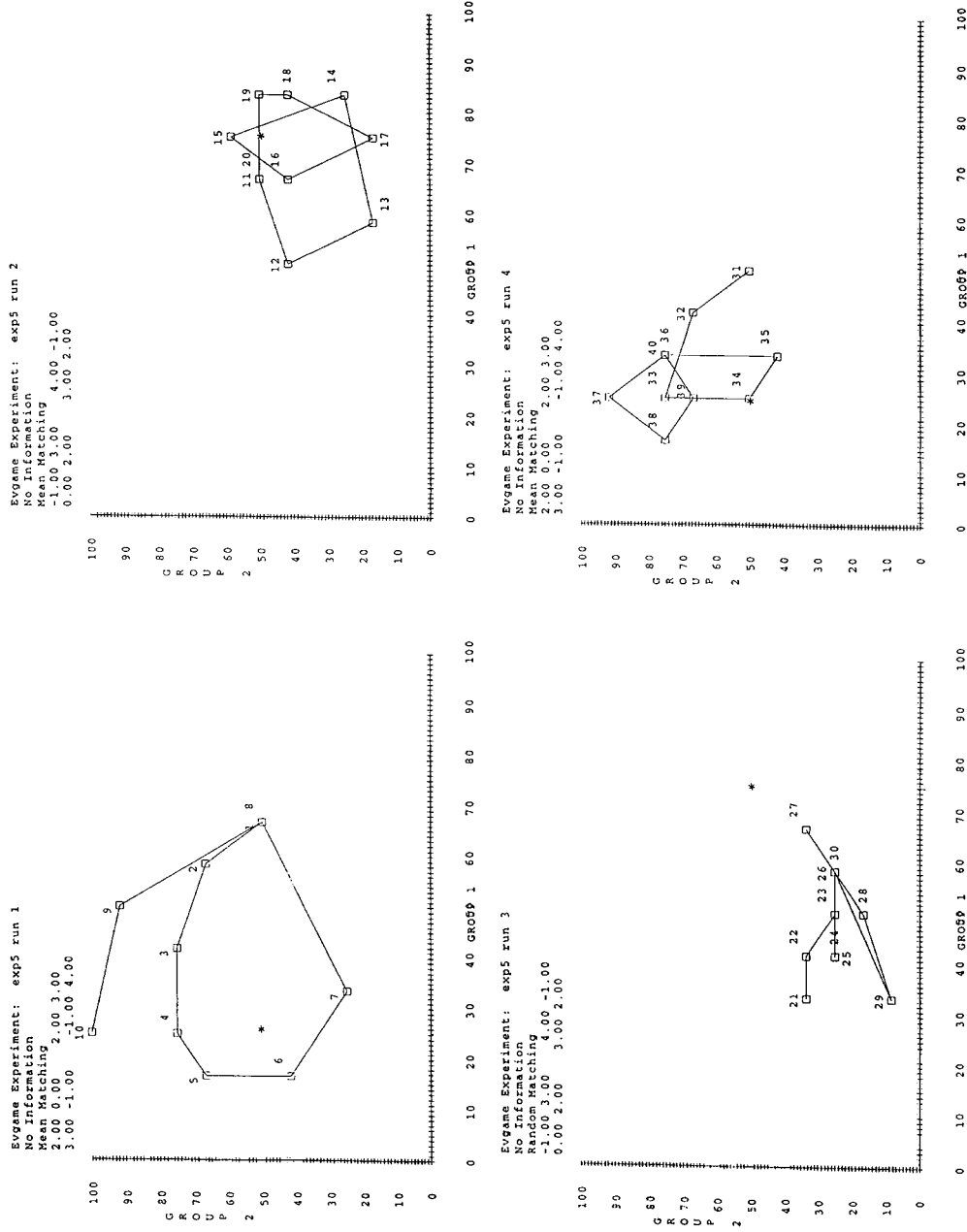


Fig. 3. Session Exp. 5, runs 1-4. Note: The equilibrium is indicated by the horizontal line.

function and that $R(s)=b+cs$ is defined for the matrix M by Eq. (1). The same set of estimating equations works for the two population (Buyer–Seller) data with $R(\hat{s}_t)$ replaced by $R_k(\hat{s}_{jt})$ for $k=1, 2$ and $j=3-k$. We allow the parameters α , β , and γ to differ across (but not within) the two populations.

Results are presented in Table 1. The maximum likelihood estimates, computed and checked using Fortran and TSP codes, indicate parameter variability across sessions and games. For the Hawk–Dove data, the ranges of the estimates are from 0.07 to 0.53 for α , from -0.65 to 0.33 for β , and from -0.45 to 1.88 for γ . Significant parameters are usually found in sessions with a larger sample size. The highly nonlinear nature of the likelihood makes it difficult to extract information from a small sample and results in large estimated standard errors.

When data are pooled across sessions, we get tighter parameter estimates. All three parameters are significantly different from zero. The steady state implied by the α estimate is $F(0.35)=0.64$, which is close to the Nash equilibrium value 0.67 for the Hawk–Dove payoff matrix. The negative β -estimate can be interpreted either as perverse behavior or as anticipatory behavior in the sense of Selten (1991). In any case, it is in the stable region as $0.20 = |\beta| < [(2\pi)^{1/2} / |c|] = 0.42$. The memory parameter γ is close to and insignificantly different from 1. This is in accord with fictitious play (or long memory) behavior as information in the past carries the same weight as the most recent one.

Parameter variability across sessions is also found in the Buyer–Seller data. The pooled data provide the most significant estimation results. The steady state corresponding to the α -estimates is (0.44, 0.47). The first component is far away from the Nash equilibrium value 0.25 while the second component is insignificantly different from the Nash value 0.5. Contrary to the Hawk–Dove case, the β estimates are all positive. Also, they are all in the stable region as $0.24 = |\beta_1\beta_2| < 2\pi / |c_1c_2| = 0.79$. The γ estimates are insignificantly different from zero.⁸

5.2. The replicator dynamics

The replicator model for the single population (Hawk–Dove) data is estimated from the equation

$$\Delta s_{t+1}/s_t = \alpha + \beta(1 - s_t)R(s_t) + \gamma d_t + \epsilon_t \quad (9)$$

The intercept α allows for the steady state (or ‘behavioral equilibrium’) to be displaced from the Nash equilibrium value.⁹ The parameter β is the adjustment speed. The variable $d_t=I(t)(1-s_t)R(s_t)$ captures the interaction of the mean matching (MM) dummy and the

⁸ Perhaps the biggest surprise in these findings is that learning parameter γ differs so much for the different games. However, CF 97 shows that this is an artifact of aggregation; the distributions of individual γ parameters are not significantly different.

⁹ Friedman and Rosenthal (1986) predict displacement towards the center of the state space. We just saw that the learning model’s estimate for Buyers suggested this sort of displacement. It should be noted that the stability analysis is essentially unaltered. For example, in the single population replicator with $\alpha \neq 0$, we get a steady state at a solution $s^\dagger \neq s^*$ to $0 = \alpha + \beta(1-s)R(s)$. Up to first order at s^\dagger , the right hand side can be rewritten as $\beta(1-s)R'(s)$ for $R'(s) = b^\dagger + cs$, where b^\dagger is $b + \alpha/[\beta(1-s^*)]$. But, by construction, $R'(s^\dagger) = 0$ and $R''(s) = c$, so that usual calculations still work at the behavioral equilibrium s^\dagger .

Table 1
Learning model parameter estimates

Panel A: Hawk–Dove data

Session	Nobs	α	β	γ
3	72	0.07 (0.38)	0.33 (0.60)	1.88 (2.64)
6	120	0.31 (0.13)	0.03 (0.16)	−0.45 (1.91)
9	120	0.30 (0.26)	−0.24 (0.23)	0.61 (0.65)
10	180	0.53 (0.15)	−0.41 (0.14)	0.81 (0.34)
15	54	0.20 (0.31)	0.32 (0.54)	0.02 (0.99)
21	81	0.17 (0.21)	−0.05 (0.17)	0.00 (1.79)
22	63	0.32 (0.24)	0.16 (0.29)	0.53 (1.82)
23	60	0.23 (0.53)	0.39 (1.04)	1.17 (1.72)
24	90	0.43 (0.20)	−0.63 (0.51)	0.42 (0.65)
25	45	0.20 (0.21)	0.62 (2.28)	0.15 (2.18)
26	90	0.28 (0.19)	−0.07 (0.35)	0.49 (2.18)
All	1131	0.35 (0.06)	−0.20 (0.08)	0.80 (0.32)

Panel B: Buyer–Seller data

Session	Nobs	Buyer			Seller		
		α	β	γ	α	β	γ
5	108	−0.27 (0.15)	0.59 (0.43)	0.12 (0.53)	0.14 (0.15)	0.38 (0.16)	0.08 (0.32)
13	44	−0.33 (0.21)	0.79 (0.99)	0.12 (1.10)	−0.01 (0.23)	0.16 (0.29)	−0.20 (1.17)
19	54	0.08 (0.19)	0.86 (0.60)	0.41 (0.48)	−0.71 (0.50)	0.56 (0.40)	0.34 (0.79)
23	60	−0.22 (0.17)	0.68 (0.65)	0.59 (0.49)	−0.51 (0.33)	0.56 (0.43)	0.55 (0.41)
27	36	0.07 (0.22)	0.76 (0.48)	−0.19 (0.42)	−0.18 (0.36)	0.35 (0.30)	0.01 (0.60)
28	36	−0.22 (0.24)	0.71 (0.70)	0.11 (0.63)	−0.16 (0.32)	0.71 (0.41)	0.20 (0.44)
29	36	−0.33 (0.33)	0.50 (0.89)	−0.22 (1.92)	0.19 (0.33)	0.54 (0.62)	0.24 (0.82)
30	36	−0.07 (0.23)	0.37 (0.62)	0.23 (0.87)	−0.19 (0.33)	0.52 (0.30)	0.21 (0.49)

Table 1
(Continued)

Panel A: Hawk–Dove data							
Session	Nobs	α	β	γ			
All	410	−0.15 (0.07)	0.74 (0.22)	0.30 (0.24)	−0.06 (0.09)	0.32 (0.09)	0.13 (0.22)

NOTE: The sessions listed in the first column are described in F96 Table 1. Nobs=($T-1$) \times number of Hawk–Dove runs in that session, where $T=10$ or 16 is the number of stages or periods per run. The first period in each run is omitted because the model includes a lagged dependent variable. The learning model given by Eq. (8) in the text is estimated using maximum likelihood. Standard errors computed from analytic second derivatives are given in parentheses. The row labelled ‘All’ gives results using data pooled over all sessions.

replicator variable, where $I(t)$ is 1 if the run at period t uses MM and is 0 otherwise. Thus, a positive γ coefficient means that the adjustment rate is faster under MM than under RP. The error term ϵ_t allows us to estimate the otherwise deterministic replicator model.

Table 2 reports the least squares estimates of Eq. (9) and their heteroskedasticity consistent standard errors. Sessions 9, 10 and 25 use only the mean matching treatment so the term γd_t is not included. The coefficient estimates generally have smaller standard errors and a lower degree of variability across individual sessions than their counterparts in Table 1. The intercept estimates in Panel A are typically small though often statistically significant. Since $\beta_s \equiv [c | s^*(1 - s^*)]^{-1} = 0.75$ for the Hawk–Dove matrix, the adjustment parameter estimates are all in the stable region (0, 1.5), with the possible exception of session 25. Indeed, with the additional exception of session 15, the estimates are in (0, 0.75), implying monotonic convergence to the equilibrium. The γ estimates are all positive (about half of them are significantly so) and the estimate from the pooled data is significantly positive. This suggests the mean matching treatment speeds the adjustment process. The overall estimated adjustment speed under MM is $0.50+0.39=0.89$, in the stable (damped) oscillation range (0.75, 1.50).

Estimates of the replicator model for the Buyer–Seller data are given in Table 2 Panels B and C. The patterns of coefficient estimates are similar to those in Panel A but with greater variability. The greater variability probably is due, at least in part, to the smaller sample sizes. But it is also worth noting that the adjusted R -squares for the Buyer–Seller data are much lower than for the simple Hawk–Dove data.

5.3. Comparison

Table 3 compares the explanatory power of the learning and replicator models using the standard root mean square error (RMSE) and mean absolute deviation (MAD) criteria. RMSE, and to a lesser extent MAD, favor models estimated by least squares over models estimated by maximum likelihood. Nonetheless we see in Panel A that the learning model has a slightly better performance. It has a lower RMSE than the replicator model in six of the 11 cases and a lower MAD in seven of the 11 cases. The maximum improvements offered by the learning model are 0.35 in RMSE and 0.15 in MAD (experiment 26) while those by the replicator dynamics specifications are 0.20 in RMSE and 0.15 in MAD. When pooled data are considered, the learning model slightly outperforms the replicator dynamics using both criteria.

Table 2
Replicator regression estimates

Panel A: Hawk–Dove data

Session	Nobs	α	β	γ	\bar{R}^2
3	72	−0.13 (0.03)	0.60 (0.12)	0.46 (0.12)	0.76
6	120	−0.08 (0.02)	0.63 (0.08)	0.20 (0.11)	0.63
9	120	−0.21 (0.04)	0.54 (0.08)		0.45
10	180	−0.10 (0.03)	0.41 (0.09)		0.32
15	54	−0.13 (0.03)	1.02 (0.13)	0.19 (0.16)	0.81
21	81	−0.20 (0.04)	0.51 (0.10)	0.37 (0.15)	0.58
22	63	−0.04 (0.02)	0.55 (0.03)	0.17 (0.14)	0.77
23	60	−0.06 (0.02)	0.61 (0.18)	0.49 (0.24)	0.62
24	90	−0.08 (0.03)	0.73 (0.19)	0.69 (0.31)	0.46
25	45	0.06 (0.03)	1.61 (0.28)		0.38
26	90	−0.00 (0.04)	0.55 (0.27)	0.14 (0.33)	0.20
All	1131	−0.08 (0.01)	0.50 (0.04)	0.39 (0.06)	0.47

Panel B: Buyer data

Session	Nobs	α	β	γ	δ	\bar{R}^2
5	108	0.06 (0.09)	0.49 (0.48)	−0.18 (0.48)		−0.2
13	44	0.15 (0.11)	0.75 (0.32)		2.17 (2.04)	0.18
19	54	0.08 (0.06)	0.83 (0.82)	1.35 (1.04)		0.36
23	60	0.19 (0.11)	0.48 (0.51)	0.72 (0.87)		0.08
27	36	0.38 (0.24)	0.92 (0.64)	3.68 (2.27)		0.39
28	36	0.63 (0.23)	2.28 (0.91)	1.27 (1.32)		0.36
29	36	0.21 (0.17)	−0.48 (0.95)	1.03 (1.05)		−0.03
30	36	0.19 (0.10)	0.94 (0.48)	0.26 (0.66)		0.16
All	410	0.23 (0.05)	0.60 (0.26)	0.23 (0.54)	0.52 (0.51)	0.12

Table 2
(Continued)

Panel C: Seller data						
Session	Nobs	α	β	γ	δ	\bar{R}^2
5	108	0.02 (0.06)	0.34 (0.21)	0.87 (0.36)		0.35
13	44	-0.03 (2.04)	0.31 (0.05)		0.37 (0.24)	0.10
19	54	-0.36 (0.14)	0.83 (0.37)	0.13 (0.32)		0.31
23	60	-0.11 (0.10)	0.53 (0.39)	-0.00 (0.40)		0.12
27	36	-0.10 (0.20)	0.33 (0.38)	0.86 (0.48)		0.43
28	36	-0.05 (0.09)	0.76 (0.33)	1.00 (0.88)		0.39
29	36	-0.05 (0.05)	0.93 (0.40)	-0.09 (0.55)		0.19
30	36	-0.15 (0.08)	0.68 (0.31)	1.14 (0.69)		0.48
All	410	-0.07 (0.51)	0.49 (0.03)	0.48 (0.14)	0.04 (0.24)	0.30

NOTE: See 'NOTE' to Table 1. The replicator dynamics model is $\Delta s_t/s_t = \alpha + \beta(1 - s_t)R(s_t) + \gamma d_t + \epsilon_t$ is fitted to data using the least squares method. The γd_t term is dropped from sessions 9, 10, and 25 in Panel A because these sessions do not use the random pairwise treatment. In session 13 and session 'ALL' of Panels B and C, the term $\delta [I(t)(1-s_t)R(s_t)]$ is included, where $I(t)$ assumes a value of 1 if history is provided to players and a value of 0 otherwise. The dummies $I(t)$ for the match and history treatments are the same for the other sessions. Heteroskedasticity consistent standard errors are given in parentheses. The row labelled 'All' gives results using observations from all the sessions.

The advantage of the learning model is much greater in the Buyer–Seller data in Panel B. The replicator model outperforms the learning model in only two instances (Buyer MAD sessions 13 and 30) and there the margin is only 0.01. In the other 30 comparisons, the learning model does at least as well and often much better – in several cases, the learning model RMSE or MAD is only half that of the replicator model.

6. Discussion

The learning model is intended to be fitted to individual players and to capture the effects of the matching and history treatments. Individual differences and treatment effects indeed are highly significant (see CF97). But to facilitate comparison to the replicator model of aggregate behavior we restricted the coefficients to be the same across individual players, ignored the treatment effects, and predicted the mean state rather than the states of the individual players. Nevertheless, the learning model slightly outperforms the replicator model for the single population (Hawk–Dove) data, even using criteria *a priori* favorable to the replicator model. The learning model vastly outperforms the replicator model on the two population (Buyer–Seller) data.

Table 3
Comparison of error estimates

Panel A: Hawk–Dove data

Session	RMSE		MAD	
	L	R	L	R
3	0.135	0.132	0.104	0.105
6	0.116	0.114	0.096	0.093
9	0.141	0.146	0.112	0.116
10	0.147	0.162	0.118	0.128
15	0.132	0.135	0.099	0.104
21	0.184	0.164	0.152	0.137
22	0.111	0.114	0.084	0.088
23	0.110	0.105	0.085	0.078
24	0.134	0.132	0.108	0.103
25	0.106	0.115	0.090	0.093
26	0.147	0.182	0.118	0.133
All	0.148	0.151	0.117	0.119

Panel B: Buyer–Seller data

Session	Buyer				Seller			
	RMSE		MAD		RMSE		MAD	
	L	R	L	R	L	R	L	R
5	0.23	0.33	0.18	0.24	0.18	0.27	0.14	0.22
13	0.14	0.15	0.12	0.11	0.12	0.15	0.09	0.13
19	0.18	0.18	0.14	0.15	0.17	0.19	0.13	0.16
23	0.19	0.28	0.16	0.22	0.19	0.25	0.15	0.21
27	0.19	0.40	0.15	0.32	0.20	0.22	0.16	0.18
28	0.16	0.37	0.12	0.28	0.14	0.26	0.11	0.22
29	0.13	0.25	0.10	0.20	0.14	0.19	0.12	0.16
30	0.19	0.21	0.17	0.16	0.16	0.20	0.14	0.16
All	0.20	0.29	0.16	0.22	0.20	0.24	0.16	0.19

NOTE: Root mean squared errors (RMSE) and mean absolute deviations (MAD) computed from the three parameter learning model are given under the column labeled 'L' and those from the replicator dynamics models are reported under 'R'. The row labelled 'All' gives results using observations from all the experiments.

We conclude that, even for explaining aggregate behavior, the learning model is more promising. It has a more consistent performance across different game structures than the replicator model, perhaps because it captures something about human decision processes. The replicator model fits the two population data rather poorly, perhaps because it presumes at best neutral stability for the mixed Nash equilibrium¹⁰ while the learning model allows for possibly stable or unstable behavior. Hence the learning model seems to provide a more robust description of the adjustment behavior.

There are a number of attractive opportunities to follow up the results reported here. First, the replicator model does reasonably well in the Hawk–Dove data, so it might be

¹⁰ Tables 3 and 4 in Friedman (1993) provide additional evidence on this point. When the replicator model (and the related 'spring' model) were expanded to allow for stabilizing own-population effects, the estimated coefficients for such effects were significant.

interesting to see if it can account for data in other single population games, for example $N \times N$ symmetric matrix games with multiple Nash equilibria or corner Nash equilibrium for $N > 2$. Second, the recent proliferation of learning models in economics (and psychology) offers many different perspectives on learning processes. Perhaps a systematic comparison of these learning models will identify a tractable and parsimonious model that can track complex human learning behavior in a variety of strategic interactions.

Acknowledgements

The experiments reported here are funded by the National Science Foundation under grant SES-9023945. We are grateful to Carl Plat for his excellent research assistance, to Thanh Lim and Tim Kolar for programming assistance, to audiences at Caltech, the University of Munich, and the 1997 ASSA meetings for their comments, and to a referee of this Journal for guidance in improving the exposition.

References

- Binmore, K., Samuelson, L., Vaughan, R., 1995. Musical chairs – modelling noisy equilibrium. *Games and Economic Behavior* 11, 1–35.
- Börgers, T., Sarin, R., 1993. Learning through reinforcement and replicator dynamics, forthcoming, *Journal of Economic Theory*.
- Brown, G., 1951. Iterated solution of games by fictitious play, In: Koopmans, T. (Ed.), *Activity Analysis of Production and Allocation*, Wiley, New York, pp. 377–380.
- Cheung, Y.-W., Friedman, D., 1997. Learning in evolutionary games: Some laboratory results. *Games and Economic Behavior* 19, 46–76.
- Crawford, V., 1995a. Adaptive dynamics in coordination games, *Econometrica* 63, 103–144.
- Crawford, V., 1995b. Theory and experiment in the analysis of strategic interaction, UCSD Economics Discussion Paper, pp. 95–37.
- Cournot, A., 1838. *Researches sur les Principes Mathematiques de la Theorie des Richesses*, English ed., In: Bacon, N. (Ed.), *Researches into the Mathematical Principles of the Theory of Wealth*, Macmillan, New York, 1897.
- Friedman, D., 1993. Sobre evolución y aprendizaje en juegos. *Cuadernos Economicos* 54, 171–196.
- Friedman, D., 1996. Equilibrium in evolutionary games: Some experimental results. *Economic Journal* 106, 1–25.
- Friedman, J.W., Rosenthal, R.W., 1986. A positive approach to non-cooperative games. *Journal of Economic Behavior and Organization* 7, 235–251.
- Fudenberg, D., Levine, D., 1998. *Learning in Games*, MIT Press, Cambridge.
- Hirsch, M.W., Smale, S., 1974. *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, New York.
- Hofbauer, J., Sigmund, K., 1988. *The Theory of Evolution and Dynamical Systems*, Cambridge University Press, Cambridge.
- Schlag, K., 1994. Why imitate, and if so, how? Exploring a Model of Social Evolution, Discussion Paper No. B-296, University of Bonn.
- Selten, R., 1991. Anticipatory learning in games, In: Selten, R. (Ed.), *Game Equilibrium Models*, chapter 3, vol. 1, Springer, New York.
- Taylor, P.D., Jonker, L.B., 1978. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40, 145–156.
- Vega-Redondo, F., 1996. *Evolution, Games and Economic Behavior*, Oxford University Press, New York.
- Weibull, J., 1995. *Evolutionary Game Theory*, MIT Press, Cambridge.