

Generality, repetition, and the role of descriptive learning models

Ido Erev^a, Ernan Haruvy^{b,*}

^a*Technion – Israel Institute of Technology, Israel*

^b*School of Management, University of Texas at Dallas, SM32, P.O. Box 830688, Richardson, TX 75083, USA*

Received 1 April 2004; received in revised form 20 June 2005

Available online 22 August 2005

Abstract

Previous studies of learning in certain classes of games appear to lead to inconsistent conclusions. Studies that focus on the observed sequential changes in behavior support models that imply high-action inertia and allow for different parameters in different games. Studies that use simulation-based analysis and focus on the prediction of behavior in new games support models that imply little-action inertia, and demonstrate the value of models that assume general parameters over certain classes of games. We show that this apparent inconsistency emerges even when analyzing a large data set with a single model. We then show that the inconsistency between the two analyses can be a product of the tendency by subjects to repeat past choices.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Learning; Inertia; Experiments

1. Introduction

Recent studies of human adjustment to economic incentives demonstrate the potential of descriptive models of learning. Learning models appear to predict well when behavior converges to equilibrium (see e.g., Erev & Roth, 1998), and to allow for predictions of behavior in situations for which equilibrium analysis does not have unique predictions (e.g., Van Huyck, Cook, & Battalio, 1997; Haruvy & Stahl, 1999; Anderson, Goeree, & Holt, 2001; Capra, Goeree, Gomez, & Holt, 2002) or when convergence to equilibrium does not occur within the experiment's timeframe (e.g., Nagel & Tang, 1998).

However, attempts to translate this potential to applications reveal that the task is not trivial. It is rarely transparent which class of learning models should be applied and the different models appear to have different implications. Seasoned researchers have expressed puzzlement over the seemingly contradicting

results different analyses yield (e.g., Feltovich, 2000; Nyarko & Schotter, 2002).

The main goal of the current paper is to try to improve our understanding of the apparent contradictions in the learning literature. The paper focuses on two related disagreements. One disagreement is regarding the value of learning models that impose general parameters across games of the type studied here. Several studies (e.g., Roth & Erev, 1995; Erev & Roth, 1998; Bereby-Meyer & Erev, 1998; Sarin & Vahid, 2000) suggest that models with general parameters over games can have good predictive ability. Yet, another stream of the literature concludes that the assumption of general parameters over games (e.g., Camerer & Ho, 1999) is counter-productive. Moreover, studies such as Stahl (1996), Cheung and Friedman (1997), Kitzis, Kelley, Berg, Massaro, and Friedman (1998), Daniel, Seale, and Rapoport (1998), Rapoport, Daniel and Seale (1998), Camerer and Ho (1998), and Busemeyer and Stout (2002) show that individuals are sufficiently different that pooling them together could result in a possible misspecification.

The second disagreement considered in this work involves the role of action inertia in games of the type

*Corresponding author. Fax: +1 972 883 6727.

E-mail address: eharuvy@utdallas.edu (E. Haruvy).

studied here. Action inertia is the tendency of decision makers to repeat past actions. Action inertia is distinguished from belief inertia in that action inertia depends on the frequency of past actions and not on past beliefs.

The basic premise of the current research is that at least a part of the inconsistencies might be resolved with a distinction between two classes of prediction tasks that can be addressed with learning models. The first class of prediction tasks involves the prediction of choice behavior in a game based on the observation of previous interaction in that same game. We refer to these tasks as *within-game* predictions. Models intended for within-game prediction tasks are useful in capturing the rules individuals use in a given game. Such models also allow for a good characterization of heterogeneity of behavior within a given game.

The second class involves the prediction of future behavior in a new game without observations on past behavior in that game. These tasks can be addressed with descriptive models by using the knowledge accumulated in the studies of other games, and using Busemeyer and Wang's (2000) generalization criterion methodology for model selection. We refer to such tasks as *new-game* prediction tasks. New game prediction tasks typically focus on commonalities and not on individual differences or differences between games. Accordingly, in the new game prediction tasks presented in this work, we do not incorporate heterogeneity. However, the lack of explicit account for heterogeneity is not a general property of new game prediction tasks. For example, one can use simulations to derive the predicted between-subjects variability for a given model (see e.g., Erev & Barron, 2005). Researchers who focus on new game predictions do not make the assertion that decision makers apply the same rules in all games or that heterogeneity is inconsequential. Rather, they seek to find commonalities over games in a given class and rules that are general enough to approximate a broad set of behaviors.

Proponents of the first approach—within game predictions—often argue that ignoring behavioral differences across games is not necessary when data is available to estimate more complex models. While this argument is correct if one seeks to investigate a given setting, it is of little practical use when one seeks to predict a new setting for which no data is available. This is particularly important when making generalizations from the laboratory to real life settings.

Under this classification, an investigation into the differences between new- and within-game predictions might clarify some of the apparent disagreements in the learning literature. In an attempt to achieve this goal, the current research compares new- and within-game predictions of behavior in a wide set of experimental conditions.

Section 2 presents the experimental conditions. We focus on 59 conditions that encompass three distinct classes of games. The analysis of a large data set is important as some of the variables considered by the leading learning models are highly correlated. Thus, reliance on small samples can lead to unreliable conclusions (see related arguments in Hopkins, 2002; Salmon, 2001; Blume, DeJong, Lowen, Neumann, & Savin, 2002). The focus on different classes of games was designed to facilitate generality.

Section 3 derives the predictions of two learning models: Experience Weighted Attraction (EWA, Camerer & Ho, 1999) and adaptive play with inertia (e.g., Cooper & Kagel, 2002). The results reveal that both models provide useful within- and new-game predictions. Yet, the models take very different forms to achieve the two tasks. Best within-game predictions are obtained with game-specific or individual-specific parameters and with parameters that point to a high degree of action inertia. Useful new-game predictions are obtained under the assumption of general parameters over games and with parameters that suggest little action inertia. Action inertia is expected to be useful in within-game prediction tasks if individuals tend to repeat actions they have chosen in the past. Our findings indicate that this is indeed the case. However, knowing that individuals repeat actions is less useful in predicting new settings with new individuals unless the researcher has near-perfect knowledge of the distribution of initial propensities in this new setting (which is typically not the case). A low degree of action inertia in new task prediction allows higher relative weight on payoff-related variables, which in the settings we investigate appear to provide useful predictions.

Note that since the homogeneous models in the two prediction tasks are structurally identical, consistency of maximum likelihood and least squares methods (Ame-miya, 1985, pp. 184–185, 270–274) implies that, *if the models correctly reflect reality*, both estimation tasks will converge in probability to the *true* parameter values as the sample size increases. However, a useful premise is that no model is perfectly specified and that models can be useful despite that fact. Section 4 presents one sufficient condition for the observed difference between within- and new-game analyses. It is shown that the difference could be a product of a psychological tendency that might seem inconsequential and undeserving of serious attention by leading learning models. Specifically, it is sufficient to assume that decision makers do not make a separate evaluation in each period. Instead, decision makers are assumed to occasionally repeat previous choices. Such repetitions have a large effect on the parameters estimated in within-game analysis, and almost no effect on the parameters estimated in new-game analysis. Thus, incorrect abstraction of these repetitions has little effect

Table 1
The 10 constant-sum no-initial information (CS-NI) games studied by Erev et al. (2002)

#	The games				Proportion of A choices ($\times 100$) by player			
	Payoff matrix				Equilibrium		Observed (trials 101–500)	
	AA	AB	BA	BB	Player 1	Player 2	Player 1	Player 2
1	77	35	8	48	49	16	70	42
2	73	74	87	20	99	79	78	38
3	63	8	1	17	23	13	52	14
4	55	75	73	60	39	45	40	51
5	5	64	93	40	47	21	26	32
6	46	54	61	23	83	67	67	42
7	89	53	82	92	22	85	20	59
8	88	38	40	55	23	26	38	14
9	40	76	91	23	65	51	58	47
10	69	5	13	33	24	33	50	20

The payoff entry ij presents the probability ($\times 100$) that Player 1 wins when she chose i and her opponent chose j . The payoff for each win was 4 cents. All 10 games were played by fixed pairs for 500 trials. Players did not receive prior information concerning the matrix, and the information after each trial was limited to the obtained payoff. Each game was played by 9 pairs.

on the value of new game predictions, but has a dramatic effect on the conclusions derived in within-game analysis. For example, when EWA is used to fit data that include repetitive behavior by subjects, the estimated parameters imply that the agents largely ignore forgone payoffs.

2. The experimental conditions

Three sets of experimental conditions are analyzed here. Two of the three sets involve “basic adaptation problems”—situations in which the economic agents do not receive prior information concerning the relative value of the possible alternatives and must base their decisions on their personal experiences. There are two reasons for this focus. The first is methodological: When agents receive prior information, the study of the adaptation process is confounded with the study of the initial tendencies (see Haruvy, 1999; Haruvy & Stahl, 2004). The second reason is our conviction that some of the most promising directions for learning research lie in the investigation of “small” repeated decisions that are made with little information and little deliberation. Whereas “big” economic decisions are likely to involve careful optimization, small decisions are likely to involve adjustment and adaptation. Though small decisions are of small consequence to the individual making them, they are potentially of tremendous importance to firms and society. In marketing, which grocery store to go to or which brand of yogurt to buy are small decisions to an individual (known as ‘low involvement’ in marketing) but are the focus of retailers. In human resource management, small decisions related to shirking, observing safety regulations, or virus protection could spell

catastrophe to the firm. Similarly in law enforcements, decisions on whether to make a full stop at a stop sign or run a red light when no one is around are small to the individual but not to society at large.¹

To facilitate comparison with other classes of games, the third set of games considered here includes constant sum games from various published studies. In these games, players receive prior information concerning the payoff rule. Notice that in order to avoid the complex interactions between reciprocation and adaptation (see e.g., Rapoport & Chammah, 1965; Bolton, 1991) we chose to avoid games in which reciprocation is feasible. The three sets of conditions are summarized below.

2.1. Random 2×2 constant sum games (the CS-NI set)

In order to demonstrate the utilization of the predictive value (PV) measure of descriptive models Erev, Roth, Slonim and Barron (2002) studied 10 randomly selected 2×2 constant sum games with a unique mixed strategy equilibrium. Table 1 presents the 10 games. The numbers in the cells of the payoff matrix represent probabilities (p_1 , p_2 , p_3 and p_4) that Player 1 will win a fixed amount w (\$0.04) if the relevant cell is selected in a given trial. The probability that Player 2 wins is 1 minus this value. For example, if both players select A on a given trial of game 1, then Player 1 wins with probability 0.77 and Player 2 wins with probability 0.23. A player who does not win the prize w earns zero

¹E.g., see driving decisions (Perry, Haruvy, & Erev, 2002), Casino gambling (Haruvy, Erev, & Sonsino, 2001), grocery shopping (Haruvy & Erev, 2004) and cheating on exams (Erev, Ingram, Raz, & Shany, 2004).

for that period. Each participant played one of the games against a fixed opponent for 500 trials. Players did not know the payoff matrix, and the only feedback after each trial was the payoff for the trial (w or 0). A player's payoff from the game was the sum of his payoffs over the 500 periods of play plus a show up fee. All choices were made anonymously via networked computers.

To derive the games, Erev et al. (2002) independently and uniformly drew the probabilities p_1 through p_4 from the set $\{0, 0.01, 0.20, \dots, 0.99, 1\}$. Games generated in this way were included in the sample if they had a unique mixed strategy equilibrium. The right-hand columns of Table 1 present the equilibrium predictions and the observed choice proportions in trials 101 to 500. Comparison of these columns shows the pattern observed by Erev and Roth (1998): In most games only one of the players moves (from uniform choice) toward the equilibrium.

2.2. Noisy 2×2 games (the $N2 \times 2$ set)

The set of 40 games to be described here was designed specifically for this study. Past learning research on games with unique mixed strategy equilibrium has been criticized on the ground that in such games the learning curves are rather flat, so the success of simple models is not surprising. Camerer (2003, p. 119) criticizes such games on the grounds that they require higher levels of mutual rationality and states that learning dynamics that move towards best response are known to spiral away from mixed equilibria (also see Crawford, 1985).² To address this criticism we chose to study random 2×2 games without any constraint on the equilibrium. To avoid the opposite problem of studying quickly converged games that cannot be used to allow robust parameter estimation, we focused on noisy games in which the entries of the payoff matrix are binary gambles of the form (Extreme, P; Moderate). To simplify the presentation, we focused on symmetric games. The following steps were taken in the derivation of each of the games:

- (1) The absolute expected payoffs of each gamble (AEPi) in points were randomly and independently selected from the interval (30, 70).
- (2) An extreme outcome was drawn for each gamble from the interval (AEPi, 100).

²Much of the learning literature, including Camerer and Ho (1999), estimated learning models on constant sum (mixed strategy) games. These games are ideal in that Maximin (a simple decision rule) coincides with Nash equilibrium. A second advantage is that dynamics take longer to converge in mixed equilibrium games. A third advantage of constant sum games is in ensuring that other-regarding behavior and cooperative behavior do not confound results since the players' interests are strictly opposed.

- (3) A moderate outcome was drawn for each gamble from the interval (0, AEPi).
- (4) The two possible outcomes were rounded to the nearest integer and the relevant probability was calculated.
- (5) The "conversion rate" from points to cents was randomly selected from the range (−0.1 to 0.1).

Table 2 presents the 40 games. The CR variable indicates the magnitude of the conversion rate. One cohort of four subjects played each of the games for 200 trials, with random matching in each trial. Random matching was used to minimize coordination and reciprocation. As in the CS-NI set, players did not know the payoff matrix, and the only feedback after each trial was their payoff for the trial. A player's payoff from the game was the sum of his payoffs over the 200 periods of play, plus a show-up fee. The show-up fee was \$12 minus the expected payoff to uniform pay in a given game. All choices were made anonymously via networked computers.

Table 2 provides the equilibrium predictions and the observed choice proportions in the last 100 trials. The results show about 78% agreements with unique pure strategy equilibrium predictions (available in 12 games). Obviously, equilibrium is less useful in predicting behavior in the 28 games with multiple equilibria.

2.3. The Constant-sum Full information games (the CS-FI set)

The CS-FI set includes nine games examined in recent learning research (see Table 3). Four of the games (1–4) were studied in Mookherjee and Sopher (1997), and reanalyzed by Camerer and Ho, (1999). Three additional games were analyzed in Erev and Roth (1998). Game 5 was originally studied in O'Neill (1987), and Games 6 and 7 in Rapoport and Boebel (1992). The final two games (8 and 9) were studied in Nyarko and Schotter (2002). In all nine conditions the players received exact prior description of the payoff matrix. The feedback after each trial included the other player's choice and the implied payoff. This information allows the calculation of the forgone payoffs. In the first eight conditions, the players were matched in pairs once. In Game 9, the players were run in cohorts of 10 subjects in each role, matched in pairs and re-matched before each trial. The right-hand columns in Table 3 present the equilibrium and the observed choice proportions in the second half of each study over participants and trials. Notice that although the games are not symmetric, the equilibrium predictions in these nine games imply identical choice probabilities by both players.

Previous analyses of these games have led researchers to seemingly different conclusions. Mookherjee and

Table 2
The 40 noisy 2×2 games

Game	CR	AA			AB			BA			BB			Pure NE	mixed NE, $Pr(A)$	Observed $P(A)$
		Ext	Mod	P_{Ext}	Ext	Mod	P_{Ext}	Ext	Mod	P_{Ext}	Ext	Mod	P_{Ext}			
1	0.052	71	23	0.17	83	45	0.13	98	58	0.03	94	47	0.04	AB, BA	0.036	0.34
2	0.062	65	31	0.12	77	43	0.41	55	52	0	83	34	0.14	AB, BA	0.49	0.46
3	0.094	57	30	0.81	94	14	0.69	52	16	0.56	74	21	0.62	AA		0.99
4	0.017	54	14	0.78	66	36	0.23	81	24	0.25	75	49	0.15	AA, BB	0.59	0.20
5	0.065	68	37	0.55	88	39	0.39	88	29	0.36	67	23	0.89	AA, BB	0.52	0.64
6	0.055	94	31	0.52	52	7	0.67	87	36	0.16	79	36	0.6	AA, BB	0.56	0.84
7	0.004	84	15	0.65	53	27	0.54	80	15	0.42	99	22	0.34	AA, BB	0.29	0.81
8	0.05	66	8	0.93	56	12	0.57	47	20	0.74	67	16	0.37	AA		0.97
9	0.025	86	14	0.76	92	25	0.54	92	31	0.33	95	26	0.43	AA		0.55
10	0.098	77	14	0.43	95	52	0.07	55	30	0.84	83	11	0.58	AB, BA	0.19	0.40
11	0.072	91	24	0.57	44	9	0.63	100	49	0.12	71	28	0.26	AA, BB	0.53	0.53
12	0.024	49	20	0.41	65	30	0.14	90	19	0.31	59	29	0.07	AB, BA	0.29	0.45
13	0.075	86	6	0.55	60	35	0	56	9	0.55	86	46	0.48	AA, BB	0.67	0.01
14	0.077	41	38	0.67	71	39	0.94	100	1	0.69	91	38	0.36	AB, BA	0.30	0.50
15	0.099	56	5	0.71	56	22	0.68	92	35	0.16	70	26	0.25	AB, BA	0.74	0.50
16	0.048	75	26	0.27	92	7	0.65	92	37	0.38	55	30	0.28	AB, BA	0.57	0.54
17	0.031	89	22	0.43	97	5	0.67	60	4	0.91	58	8	0.92	AB, BA	0.75	0.60
18	0.038	-85	-43	0.45	-91	-63	0.07	-95	-5	0.33	-63	-36	0.04	BB		0.02
19	0.011	-54	-25	0.62	-82	-2	0.4	-98	-48	0.24	-80	-14	0.41	AA		0.82
20	0.056	-91	-27	0.2	-86	-26	0.6	-87	-53	0.03	-72	-34	0.45	AA, BB	0.43	0.87
21	0.02	-98	-60	0.24	-98	-14	0.58	-74	-17	0.42	-55	-4	0.88	BB		0.26
22	0.088	-96	-28	0.28	-85	-2	0.61	-71	-19	0.27	-96	-35	0.46	AB, BA	0.43	0.45
23	0.021	-88	-3	0.4	-83	-6	0.69	-75	-34	0.51	-64	-8	0.59	AA, BB	0.53	0.20
24	0.083	-98	-6	0.67	-51	-41	0	-88	-54	0.29	-51	-17	0.76	AB, BA	0.33	0.56
25	0.004	-49	-29	0.5	-69	-23	0.2	-92	-31	0.52	-81	-34	0.57	AA		0.94
26	0.087	-85	-43	0.17	-92	-36	0.05	-76	-56	0.35	-78	-11	0.61	AA		0.86
27	0.085	-81	-25	0.3	-84	-5	0.8	-90	-31	0.61	-95	-66	0.03	AA, BB	0.52	0.98
28	0.096	-68	-54	0.64	-75	-26	0.47	-79	-48	0.48	-95	-31	0.16	BB		0.61
29	0.092	-72	-23	0.65	-79	-53	0.04	-48	-33	0.33	-93	-43	0.4	AB, BA	0.35	0.48
30	0.014	-60	-50	0.1	-92	-27	0.38	-78	-46	0.16	-81	-6	0.45	AA, BB	0.99	0.19
31	0.009	-95	-42	0.21	-97	-33	0.19	-78	-34	0.18	-70	-31	0.67	AB, BA	0.52	0.53
32	0.03	-98	-45	0.42	-85	-23	0.18	-70	-26	0.8	-87	-25	0.37	AB, BA	0.69	0.62
33	0.087	68	13	0.95	59	24	0.97	98	12	0.53	72	42	0.6	AA, BB	0.22	0.38
34	0.004	85	36	0.24	63	28	0.17	77	37	0.58	45	8	0.73	BB		0.36
35	0.068	-96	-55	0.29	-54	-40	0.86	-80	-49	0.55	-85	-63	0.23	AB, BA	0.95	0.73
36	0.068	-91	-53	0.05	-62	-10	0.69	-80	-28	0.23	-75	-30	0.73	AB, BA	0.53	0.47
37	0.081	76	9	0.81	96	41	0.31	85	19	0.36	68	42	0.35	AA		0.71
38	0.036	85	36	0.55	91	11	0.45	100	39	0.11	86	44	0.57	AA, BB	0.55	0.35
39	0.049	-95	-43	0.02	-79	-22	0.47	-97	-11	0.59	-89	-46	0.37	AA		0.72
40	0.056	-97	-48	0.1	-98	-32	0.5	-73	-54	0.74	-58	-43	0.4	AA, BB	0.51	0.95

The values in entry ij presents the gamble (Ext, P_{Ext} ; Mod) played by the player when she chose i and her opponent chose j . For example, expected payoffs to row player in game 1 are (AA = 31.16; AB = 49.94; BA = 59.2; BB = 48.88). Each game was played by a single cohort of 4 randomly matched players for 200 trials. The information after each trial included the obtained payoff, the forgone payoff and the opponent selection. The last column denotes the observed proportion of A choices in the last 100 trials of each game.

Sopher's (1997) analysis of Games 1–4 supports belief learning (fictitious play). Camerer and Ho's (1999) analysis of the same games supports EWA. O'Neill's analysis of Game 5 supports the minimax predictions, but Brown and Rosenthal (1990) and Shachat (2001) qualified this assertion. Erev and Roth's (1998) analyses of Games 5–7 support reinforcement learning. Finally, Nyarko and Schotter's (2002) analysis of Games 8 and 9 support belief learning and EWA (depending on the information used to make the predictions).

3. Comparison of new- and within-game predictions

3.1. Two models

In order to clarify the differences between new- and within-game predictions, the current analysis focuses on two learning models that were designed to allow for both types of prediction. To better follow the discussion of the models, it is useful to first define action inertia. *Action inertia* is defined to occur when a mere selection

Table 3
The constant sum full information (CS-FI) games

Payoff matrix					Game parameters		Equilibrium and observed proportions in the second half								
Games 1 & 2 (Mookherjee & Sopher, 1997)															
	A2	B2	C2	D2		T #S		A	B	C	D				
A1	W	L	L	W	1: W = 5, L = 0	40 10 × 2	Row	0.30	0.23	0.43	0.05				
B1	L	L	W	W			Col	0.44	0.30	0.26	0.01				
C1	L	W	#	#	2: W = 10, L = 0	40 10 × 2	Row	0.38	0.13	0.40	0.10				
D1	L	L	&	W			Col	0.29	0.33	0.35	0.03				
							Eqm.	0.375	0.25	0.375	0				
Games 3 & 4 (Mookherjee & Sopher, 1997)															
	A2	B2	C2	D2	E2	F2		A	B	C	D	E	F		
A1	W	L	L	L	L	W	3: W = 5, L = 0	40 10 × 2	Row	0.42	0.20	0.04	0.10	0.25	0.01
B1	L	L	W	W	W	W			Col	0.39	0.33	0.08	0.13	0.08	0
C1	L	W	L	L	W	L	4: W = 10, L = 0	40 10 × 2	Row	0.41	0.22	0.05	0.06	0.25	0.03
D1	L	W	W	L	L	L			Col	0.32	0.23	0.19	0.13	0.12	0.02
E1	L	W	L	W	L	W			Eqm.	0.375	0.25	0.125	0.125	0.125	0
F1	L	L	W	L	W	W									
Game 5 (O'Neill, 1987)															
	A2	B2	C2	D2				A	B	C	D				
A1	W	L	L	L	5: W = 5, L = -5	105 25 × 2	Row	0.36	0.24	0.21	0.19				
B1	L	L	W	W			Col	0.44	0.24	0.17	0.16				
C1	L	W	L	W			Eqm.	0.40	0.20	0.20	0.20				
D1	L	W	W	L											
Games 6 & 7 (Rapoport & Boebel, 1992)															
	A2	B2	C2	D2	E2			A	B	C	D	E			
A1	W	L	L	L	L	6: W = 10, L = -6	120 10 × 2	Row	0.29	0.27	0.16	0.10	0.17		
B1	L	L	W	W	W			Col	0.33	0.18	0.19	0.14	0.16		
C1	L	W	L	L	W	7: W = 15, L = -1	120 10 × 2	Row	0.29	0.35	0.12	0.11	0.14		
D1	L	W	L	W	L			Col	0.36	0.20	0.21	0.11	0.13		
E1	L	W	W	L	L			Eqm.	0.375	0.25	0.125	0.125	0.125		
Games 8 & 9 (Nyarko & Schotter, 2002)															
								A	B						
						8: fixed match	60 13 × 2	Row	0.47	0.53					
								Col	0.38	0.62					
A1	6	3				9: rand match	60 30	Row	0.44	0.56					
B1	3	5						Col	0.32	0.68					
								Eqm.	0.40	0.60					

The payoff matrix presents row's payoff. The constant sum is $L + W$ in Games 1–7, and 10 in games 8 and 9. # and & are the gambles ($W, 1/3; 0$) and ($W, 2/3; 0$) respectively. The payoffs are in rupees in Games 1–4 (that were run in India), and US cents in the other games 5, 8 and 9. In Games 6 and 7 subjects were paid \$1 per point in three randomly selected trials. The conversion rate from rupees to cents was about 4 cents per rupee. Yet, the purchasing power (at least rent-wise) was about 100 cent per rupee.

of a particular action increases the probability of repeated choices of this action even when the obtained outcomes are *disappointing* relative to the payoffs that could be obtained from a selection of a different option. As we will show in the next section, the assumption that reinforcements are summed (e.g., Roth & Erev, 1995; Camerer & Ho, 1999) implies a high degree of action inertia. Fictitious play models (e.g., Fudenberg & Levine, 1998) and reinforcement average models (e.g., Erev, Bereby-Meyer, & Roth, 1999) imply low rates of action inertia. Some models (e.g., Cooper & Kagel, 2002) explicitly assume action inertia. Note that action inertia as defined here is related to but distinct from inertia in beliefs, reinforcements, attractions or propensities. In fact, action inertia is closer to the tendency to

prefer the status quo (Samuelson & Zeckhauser, 1988). Models that imply high rates of action inertia were supported in some analyses (e.g., Roth & Erev, 1995; Chen & Tang, 1998; Camerer & Ho, 1999; Cooper & Kagel, 2002). However, other analyses support models that imply low rates of action inertia (see e.g., Erev et al., 1999; Haruy & Erev, 2004).

3.1.1. Experience weighted attraction (EWA)

Camerer and Ho's (1999) EWA model is presented as a hybrid approach between belief and reinforcement learning, where the two schools are described as differing on the weight given by human players to foregone payoffs in updating propensities. As noted by Camerer and Ho (1999), belief learning does not

distinguish foregone from realized payoffs in the updating of propensities, whereas decision makers in reinforcement learning models do not pay any attention to foregone payoffs. The idea behind the EWA model is simple and intuitive: decision makers evaluate the performance of each possible action in the last period and update their propensities to use each action accordingly. However, the action actually chosen by each decision maker receives greater attention in the evaluation process. Hence, actions are reinforced according to past performance, but actions actually selected receive some additional reinforcement.

The choice probabilities are expressed as logistic functions of the propensity, or attraction, of each strategy, A_{njt} , where n indexes the individual, j the strategy, and t the period of play:

$$P_{njt} = \frac{\exp(\lambda A_{njt})}{\sum_{j \in J} \exp(\lambda A_{njt})}, \quad (1)$$

where λ can be thought of as a precision parameter, or the sensitivity of players to attractions. The attractions are updated according to the following two dynamics:

$$A_{njt+1} = \frac{\phi N_{nt} A_{njt} + [\delta + (1 - \delta)\chi_{njt}] u_{njt}}{N_{nt+1}} \quad (2)$$

where u_{njt} is the actual payoff, realized or foregone, to action j in period t , and

$$N_{nt+1} = \phi(1 - \kappa)N_{nt} + 1. \quad (3)$$

The function χ_{njt} , is a characteristic function that takes the value of 1 if player n played action j in period t and 0 otherwise, and δ is an ‘imagination parameter’ which determines the ability of a player to consider payoffs that could have resulted were he to have selected a different action. There are two decay parameters, ϕ and κ .

The initial value N_{n1} is a free parameter referred to as $N(0)$. In the original EWA formulation, a free parameter A_{j0} was estimated for each possible action j . To avoid an excessive number of free parameters we assume that A_{nj1} equals the role-specific expected payoffs from random choice.

To highlight the conditions under which EWA entails action inertia, note that in EWA the change in the attraction between trial t and $t + 1$ for the selected action k is

$$A_{nkt+1} - A_{nkt} = \frac{(\phi N_{nt} - N_{nt+1})A_{nkt} + u_{nkt}}{N_{nt+1}}. \quad (4)$$

For an unchosen action $j \neq k$, the change is

$$A_{njt+1} - A_{njt} = \frac{(\phi N_{nt} - N_{nt+1})A_{njt} + \delta u_{njt}}{N_{nt+1}}. \quad (5)$$

An increase in the ratio of probabilities P_{nkt}/P_{njt} from trial t to trial $t + 1$ occurs if and only if $A_{nkt+1} - A_{njt+1} > A_{nkt} - A_{njt}$. From the two equations

above, this condition can be rewritten as

$$(\phi \kappa N_{nt} - 1)(A_{nkt} - A_{njt}) + u_{nkt} - \delta u_{njt} > 0. \quad (6)$$

Recall that action inertia was defined to occur when a selection of a particular action increases the probability of repeating this action regardless of outcome. Now that we have defined our notation, we can state this definition more formally. Action inertia occurs when a selection of action k results in $P_{nkt} > P_{njt}$ when $u_{njt} > u_{nkt}$.

Consider the case of disappointing yet positive payoffs: $u_{njt} > u_{nkt} > 0$. When $A_{nkt} = A_{njt}$, action inertia (in the comparison of j and k) occurs if $u_{njt} > u_{nkt} > \delta u_{njt}$. Thus, the likelihood of action inertia increases as δ decreases. In addition, the likelihood of action inertia increases with $(\phi \kappa N_{nt} - 1)(A_{nkt} - A_{njt})$. Under the assumption of uniform choice at trial t , the latter effect is cancelled out.

Interestingly, when the payoffs are negative, the above example entails little or no action inertia (because $u_{njt} < \delta u_{njt}$). Indeed, the opposite pattern can be predicted. A decrease in the probability of repeated choice (of k) is likely when $u_{njt} < u_{nkt} < \delta u_{njt} < 0$.

We caution that the co-movement of delta and action inertia was derived only for the case where $A_{nkt} = A_{njt}$. For the general case, A_{nkt} and A_{njt} are themselves functions of δ , which somewhat weakens the argument because the relationship between δ and action inertia may be non-monotonic. That is, in EWA, $\delta < 1$ is a necessary but not sufficient condition for action inertia.

3.1.2. Adaptive play with inertia (API)

To allow direct examination of action inertia we chose to study a quantification of the idea of adaptive play with inertia, referred to as API. The choice rule is stochastic and beliefs are updated with depreciation of past beliefs by a constant. Following Erev et al. (1999) the model involves normalization (division by a measure of payoff variability) of the propensities. This normalization was introduced to capture the effect of payoff variability, explained in Section 2.1. Lastly, the model explicitly captures action inertia. Action inertia is modeled by the assumption that the propensity to select an action is a weighted average of past payoffs and a measure of past relative frequency of choice by the individual player (Cooper and Kagel, 2002, use a similar action inertia assumption).

As in EWA, API applies the logit probabilistic choice rule:

$$P_{njt} = \frac{\exp(\lambda A_{njt})}{\sum_{j \in J} \exp(\lambda A_{njt})}, \quad (7)$$

where P_{njt} is the probability of player n choosing action j in period t and A is the propensity towards action j , also known as ‘attraction’ in EWA. The term λ is a precision or sensitivity value. Propensities are updated as a weighted average of rescaled expected payoffs and

relative frequency of past choices:

$$A_{njt} = (1 - \omega) \frac{\bar{U}_{njt}}{S_{nt}} + (\omega)\pi_{njt}. \tag{8}$$

π_{njt} is the weighted relative frequency of past choices of action j by period t . S_{nt} is a normalization factor, to be described below. \bar{U}_{njt} is the expected payoff from action j in period t , based on weighted past observations on payoffs, with equal weight given to foregone and realized payoffs. That is,

$$\bar{U}_{njt} = (1 - \phi_1)\bar{U}_{njt-1} + \phi_1 u_{njt-1}, \tag{9}$$

where u_{njt-1} is the actual payoff, realized or foregone, to action j in period $t-1$. As in EWA the initial value \bar{U}_{nj1} is the expected payoff from random play. Past frequency of play is updated in a similar manner:

$$N_{njt} = (1 - \phi_2)N_{njt-1} + \chi_{njt-1}, \tag{10}$$

where χ_{njt} , as in EWA, is a characteristic function that takes the value of 1 if player n played action j in period t and 0 otherwise, and for $t > 1$.

$$\pi_{njt} = \frac{N_{njt}}{\sum_{k \in J} N_{nkt}}. \tag{11}$$

The relevant initial values are $\pi_{nj1} = 1/J$ and $N_{nj0} = 0$.

Following Busemeyer (1985) and Erev et al. (1999), the normalization factor in Eq. (8), S_{nt} , is a measure of observed payoff variability. To define S_{nt} , let y_{nt} be the realized payoff and x_{nt} the foregone payoff to player n (from the most attractive alternative option) at time t . Then:

$$a_{nt} = \frac{(t - 1)a_{nt-1} + y_{nt-1}}{t}, \tag{12}$$

$$v_{nt} = \min\{|y_{nt} - x_{nt}|, |a_{nt} - y_{nt}|\}, \tag{13}$$

and

$$S_{nt+1} = \frac{tS_{nt} + v_{nt}}{t + 1}. \tag{14}$$

Eq. (12) is the updating of average payoffs, a_{nt} , such that each new observation receives the same weight as all previous observations. Eq. (13) is the absolute deviation in the most recent period from the average payoffs or from the foregone payoffs (whichever is smaller). Eq. (14) gives the updating to the measure of payoff variability in the game. S_{n1} is initialized at the expected absolute difference between a_{n1} and the obtained payoffs from random choice.

To summarize, the model has four parameters: ω —the weight on inertia, λ —payoff sensitivity parameter, ϕ_1 —the payoff averaging parameter, and ϕ_2 —the inertia updating parameter.

We stress that the main purpose of API is not to propose an alternative to EWA but rather to allow a more direct interpretation of action inertia. In the API

model, action inertia is captured explicitly, whereas in EWA it is captured implicitly through δ .

3.2. Within-game prediction

The common goal of within-game predictions is to use all the information observed prior to trial $t + 1$ in predicting choice behavior at that time. Typically accuracy is measured using the Log-Likelihood criterion, which for a given observation would be

$$LL = \sum_{k=1}^m P_{k,obs} \ln(P_{k,pred}) \tag{15}$$

where $P_{k,obs}$ takes on the value 1 if choice k was selected and 0 otherwise. $P_{k,pred}$ is the probability of choice k predicted by the model. The final log-likelihood of a condition is calculated by adding up the log-likelihood over observations.

3.2.1. Parameter generalization over games

In order to test the assumption of general parameters over games we fitted each of the two models to each of the conditions in each of the four sets using maximum likelihood. Table 4a presents the LL scores in each set of games under three constraints: (1) One set of parameters within each set, (2) one set of parameters to each game and player role, and (3) one set of parameters for each individual. Table 4a shows large differences between the LL scores under the three constraints in all six comparisons (two models by three data sets). A likelihood ratio test reveals that these differences imply that the assertion of one set of parameters per data set can be rejected in favor of the weaker constraint of role specific parameters ($p < 0.0001$ in all cases). Table 4a shows that the constraint of role specific parameter can be rejected in favor of the weaker constraint of “individual specific parameters”.³

Notice that the results show a large action inertia component. In EWA, where δ is potentially indicative of action inertia, the mean δ is below 0.5. In API, where ω is the action inertia, the mean ω is above 0.5. The estimated inertia component decreases with the move from “set specific” parameters to “role and game specific parameters”. Allowing “individual specific” parameters reduces the estimated inertia value further. This trend is consistent with the assertion that at least

³Notice that both models assume sensitivity to payoff magnitude. To derive these models’ predictions, we first converted the payoffs to their US dollar equivalence. In particular, payments in US experiments were adjusted for inflation, and the payments in the Mookherjee and Sopher’s study were converted at an exchange rate of 1 rupee = 4 cents. In a second analysis, the payoffs in Mookherjee and Sopher’s study were converted based on their purchasing power in India at the time, and 1 rupee was assumed to equal 100 cents. Since the two analyses yield qualitatively similar comparisons, only the first is reported.

Table 4
Within-game predictions by data set and estimation constraint

	EWA						API						
	N(0)	λ	δ	ϕ	κ	LL	MSD	λ	ω	ϕ_1	ϕ_2	LL	MSD
<i>(a) When the parameters are estimated and fit is computed based on all trials</i>													
With set specific parameters													
CS-NI	5.575	3.00	0	0.567	0	-45,797	0.171	2.440	0.767	0.845	0.121	-45,315	0.169
N2 × 2	0.724	7.118	1	0.970	0.876	-18,392	0.199	2.737	0.856	0.553	0.079	-14,470	0.151
CS-FI	0.055	0.006	0.037	0.957	1	-21,380	0.173	2.642	0.745	0.059	0.012	-20,807	0.167
With game and role specific parameters (mean reported)													
CS-NI	105.7	2.276	0.026	0.576	0.420	-44,951	0.168	2.338	0.715	0.805	0.268	-43,476	0.162
N2 × 2	14.61	59.9	0.60	0.680	0.697	-15,249	0.160	3.713	0.642	0.359	0.264	-13,931	0.145
CS-FI	16.25	0.107	0.338	0.951	0.715	-20,766	0.165	4.644	0.465	0.065	0.087	-20,560	0.164
With game and individual specific parameters (mean reported)													
CS-NI	91.49	3.595	0.131	0.484	0.376	-39,897	0.150	3.283	0.558	0.666	0.354	-38,381	0.143
N2 × 2	16.46	78.62	0.638	0.519	0.701	-13,802	0.145	5.713	0.575	0.263	0.419	-12,495	0.130
CS-FI	25.49	0.399	0.701	0.761	0.606	-20,030	0.157	2.238	0.666	0.058	0.062	-19,874	0.156
<i>(b) When the parameters are estimated based on the first part (first r trials with r = Min(T/2, 100)) of the study and fit is measured on the second part (trials r+1 to the end)</i>													
With set specific parameters													
CS-NI	3.893	2.688	0	0.555	0	-36,267	0.168	2.033	0.711	0.790	0.158	-35,955	0.166
N2 × 2	0	7.390	1	0.963	1	-8884	0.189	2.490	0.839	0.483	0.105	-6769	0.139
CS-FI	0.868	0.007	0.066	0.960	1	-10,832	0.175	4.711	0.575	0.033	0.0001	-10,525	0.169
With game and role specific parameters (mean reported)													
CS-NI	51.6	2.166	0.050	0.546	0.311	-35,990	0.168	2.051	0.629	0.767	0.186	-35,258	0.163
N2 × 2	18.10	68.51	0.61	0.60	0.62	-7658	0.157	5.306	0.483	0.206	0.352	-7002	0.143
CS-FI	22.99	0.129	0.371	0.947	0.587	-10,784	0.173	4.711	0.343	0.066	0.092	-10,591	0.170
With game and individual specific parameters (mean reported)													
CS-NI	63.67	2.526	0.223	0.533	0.575	-65,534	0.172	3.546	0.402	0.633	0.326	-38,612	0.165
N2 × 2	16.37	76.45	0.633	0.553	0.737	-10,335	0.165	6.202	0.587	0.263	0.391	-7350	0.138
CS-FI	23.45	0.635	0.693	0.753	0.511	-11,819	0.183	4.581	0.264	0.173	0.353	-10,885	0.178

Table 5
API without division by variance, based on all trials

	API—without division by variance						
	λ	ω	ϕ_1	ϕ_2	LL	MSD	
<i>With set specific parameters</i>							
CS-NI	3.300	0.566	0.851	0.115	-44,976	0.168	
N2 × 2	23.89	0.895	1.00	0.086	-14,808	0.155	
CS-FI	2.055	0.017	0.060	0.0001	-20,877	0.169	
<i>With game and individual specific parameters (mean reported)</i>							
CS-NI	3.177	0.525	0.814	0.265	-43,447	0.162	
N2 × 2	13.18	0.802	0.841	0.190	-14,376	0.150	
CS-FI	3.829	0.369	0.069	0.047	-20,556	0.164	
<i>With game and individual specific parameters (mean reported)</i>							
CS-NI	4.801	0.387	0.679	0.327	-38,357	0.143	
N2 × 2	11.53	0.717	0.626	0.334	-13,550	0.141	
CS-FI	3.188	0.647	0.154	0.292	-19,847	0.155	

part of the estimated inertia is a result of using incorrect models. We return to this assertion in Section 4.

In addition, the results show a clear advantage of API over EWA. Indeed API outperforms EWA in all four

sets of conditions and different levels of generality. This advantage is noteworthy as EWA has more parameters.

We also examine a variation on API (Table 5), without the division by observed payoff variability. This

variation allows us to examine whether the division by payoff variability is responsible for the better fit of API relative to EWA in within game tasks. The results show that in within game tasks, this simpler API model does as well or better than the original.

3.2.2. Predicting future choices

Table 4b presents the within-game predictive scores of the two models in predicting the choice behavior “later” in the experiment when the parameters are estimated based on the earlier part of the experiment.⁴ In experiments with more than 200 trials, “later” is defined as all the trials after trial 100. For experiments with less than 200 trials, “later” is defined as the second half of the experiment. The results show the pattern also observed in Table 4a: Parameter estimates pointing at high action inertia, an advantage to role specific parameters, and an advantage of API over EWA.

In addition to the LL scores, Table 4 presents the Mean Squared Deviation (MSD) scores between the predictions and the data. Like the LL scores, the MSD scores were calculated at the level of the individual observation. In the case of m choices, this score is

$$\text{MSD} = \frac{1}{m} \sum_{k=1}^m [(1 - P_{k,\text{obs}})P_{k,\text{pred}}^2 + P_{k,\text{obs}}(1 - P_{k,\text{pred}})^2], \quad (16)$$

where $P_{k,\text{obs}}$ takes on the value 1 if choice k was selected and 0 otherwise. $P_{k,\text{pred}}$ is the probability of choice k predicted by the model. The final MSD of a condition is calculated by adding up the MSDs over observations and dividing by the number of observations. The MSD scores show the same picture presented by the LL scores.

3.3. New-game predictions

Recall that the derivation of new-game predictions is only possible under the assumption of general parameters across games (because it involves the prediction of behavior in a new game with the parameters estimated in other games). The clear rejection of this assumption in previous research and Section 3.2.1 implies that it may not be possible to derive accurate new-game predictions with the current models. Nevertheless, it is possible that in certain classes of games the inaccuracy of new-game predictions based on learning models will be smaller than the inaccuracy of new-game predictions derived using the popular equilibrium

models. Indeed, this idea is one of the main contributions of Erev and Roth (1998). They analyzed 12 matrix games with unique mixed strategy solutions and showed that the new-game predictions of reinforcement and belief-based learning models in that set were more accurate than the equilibrium predictions. The current analysis will examine if this advantage of learning models holds in the wider set of 59 games studied here.

In order to clarify the idea of new-game predictions, the current analysis focuses on the predictions of the aggregated choice proportions presented in Tables 1–3. We compare the models’ predictions to Nash equilibrium predictions. The first row of results in Table 6 presents the MSD scores for the equilibrium predictions.⁵

In the case of multiple equilibria (28 of the 40 games in the $N2 \times 2$ set), we computed the MSD score of the mixed strategy equilibrium. This resolution of the multiplicity led to lower MSD score than other possible resolutions. The minimal MSD scores (The MSD scores of accurate aggregate predictions) are presented in the second row.

In the first stage of the analysis we estimated the parameters of the two models in each of the four classes of conditions. A computer simulation-based method was used. That is, for each class of conditions and each model, we ran computer simulations in which virtual players that behave according to the model’s assumptions participated in a virtual replication of the class’ experimental conditions. A grid search estimation method was used. A wide set of possible values for the different parameters were considered. Then one hundred simulations were run in each of the conditions under each set of parameter values. The simulations were run for the same number of trials as the experimental players (e.g., 500 trials in the CS-NI data set). The initial values were computed based on the payoff matrix. The following steps were taken in trial t of each simulation:

- (1) The virtual players chose an action.
- (2) The payoff of the observed actions was determined based on the problem payoff rule.
- (3) The relevant statistic was recorded.
- (4) The propensities and the other values assumed to affect future choices were updated.

Each set of simulations resulted in a predicted choice proportion in each of the conditions under each set of parameters. The predictions were summarized with the same statistic used to summarize the experimental results.

⁴We use this method because it is a common approach for obtaining external validity for a model in the research we consider (e.g., Camerer & Ho, 1999). It is based on the assumption that the learning parameters do not change during the learning process. More recent research (Chen & Khoroshilov, 2003; Chen & Gazzale, 2004) addresses the limitations of this assumption.

⁵We focus on MSD rather than on likelihood scores because even a slight deviation from the pure equilibrium prediction would lead to a log likelihood of negative infinity.

Table 6
New-game predictions

Model, data set used for estimation, and parameters						Prediction error (MSD)		
						CS-NI	N-2 × 2	CS-FI
Equilibrium prediction						0.2512	0.2597	0.1749
Minimal (the score of an accurate model ¹)						0.2134	0.1802	0.1718
EWA	$N(0)$	λ	δ	ϕ	κ			
CS-NI	0.0001	5	0.8	0.7	0.01	(0.2174)	0.2100	0.1748
N2 × 2	0.0001	60	1	0.3	0.1	0.2230	(0.1990)	0.1785
CS-FI	0.0001	1	1	0.7	0.3	0.2197	0.2147	(0.1741)
API	λ	ω	ϕ_1	ϕ_2				
CS-NI	1.5	0.1	0.1	0.1		(0.2174)	0.2065	0.1748
N-2 × 2	12	0	0.7	–		0.2248	(0.1983)	0.1772
CS-FI	4	0.1	0.05	0.4		0.2258	0.2184	(0.1733)

The left-hand columns show the estimated parameters in the four sets of games. The right-hand columns show the MSD score of the models with these parameters in each set (the ij th cell uses parameter estimates from the i th model to predict the j th model). Thus, the diagonal shows the fitted values, and the off-diagonal scores show the new game predictions.

¹An “accurate model” is a model with predicted probabilities equal to the observed proportions of play.

In the second stage of the analysis, the parameters estimated in each class of games were used to predict behavior in the other three sets of games. Table 6 summarizes the main results. It shows a clear advantage of the two learning models over equilibrium.⁶ These results demonstrate the value of the assumption of general parameters across tasks.

An inspection of the estimated parameters reveals very little action inertia: In API the inertia parameter ω is close to zero in all three cases. In EWA the estimated δ value is close to 1.⁷ When $\delta = 1$ in EWA, no action inertia exists.

Finally, note that the within-game MSD scores (presented in Table 4b) are much lower than the new-game scores (Table 6). Indeed, they are lower than the best (minimal) MSD scores that can be obtained in new-game predictions. This result is hardly surprising. It is common sense that when deriving predictions it is better to use all the available information. Thus, new-game predictions should not be used when the game is not new and sufficient data to derive robust within-game predictions is available.

⁶In additional analysis, we examined the one-parameter quantal response equilibrium concept proposed by McKelvey and Palfrey (1995), and a version of API without normalization (division by payoff variability). Under the constraint of general parameter over sets of games, the quantal response model provides a better fit than the Nash equilibrium, but it is outperformed by the learning models. The new game predictions of API without normalization are outperformed by the new game predictions of API with normalization. It seems that the normalization feature of API is useful for new game predictions.

⁷The two models were also estimated and evaluated using a likelihood criterion. This analysis shows the results obtained with the MSD criterion.

3.4. The relationship between within- and new-game predictions

Under one explanation of the difference between within-game and new-game analyses, the difference is a consequence of the fact that within-game analysis (e.g., the likelihood method used in Section 3.2) uses more data, and for that reason is more accurate. In order to evaluate this explanation we re-derived the predictions of API and EWA for the four sets of games using the ML parameters estimated in Section 3.2 (with set-specific parameters). The results show that the within-game parameters tend to impair the simulation-based predictions. For example, with the ML estimated parameters for the N2 × 2 set, simulation of EWA leads to an MSD score of 0.2246. This score is larger than the score of the simulation based prediction of EWA with the parameters estimated on each of the other data sets.

3.5. The possibility of individual and task specific models

Recall that one obvious contributor to the large action inertia parameters estimated in the within-game analysis is the fact that the models studied here are abstractions and therefore imprecise. As noted above, decreasing potential misspecification by allowing for game, role and individual parameters may decrease the estimated action inertia. Removing additional constraints might reduce the estimated action inertia further. Whereas such constraints are necessary to address new-game predictions, since no data exists on the new game or the decision makers, they can be

relaxed in the derivation of within-game predictions, where such data can be collected.⁸

4. The role of repetition and sufficient conditions

Section 3.4 demonstrated that the current results could not be explained with the assertion that the within-game estimates are simply more accurate. Rather, the difference is likely to reflect the fact that the models considered here are useful but imprecise approximation. As a result, estimates that fit one set of statistics may fail to capture another set of statistics.

Our demonstration so far has shown that parameters estimated in new-game analysis may be more useful in predicting behavior in new data sets, but we have yet to show why that is the case. To do so, it is necessary to specify which of the implicit imprecise simplifications assumed by the models considered here biases the parameters estimates in within-game analysis more than they bias the parameters estimated in new-game analysis. The current section takes one step in this direction. It highlights one common simplification assumption that can account for the current results.

The assumption considered here involves the timing of decisions. The current models assume that agents in multi-trial experiments make a new decision in each trial. This assumption seems benign, but it ignores the reasonable possibility that in some trials the participants simply follow the strategy selected in previous trials rather than selecting a new strategy. Repetition of past decisions can save effort, and in some cases can be justified based on reasonable considerations (for example, if the agent believes that the payoff from one of the options increases with repeated choices).

To clarify the implications of repetition it is convenient to consider a numerical example. Consider a group of homogeneous agents that behave according to a simplified variant of the EWA model. With one exception, the agents behave according to EWA with the parameters $N(0) = 0.0001$, $\delta = 1$, $\phi = 0.5$, $\kappa = 0.5$, and $\lambda = 0.05$ (in this illustration we do not multiply payoffs by the exchange rate). The exception is that every round (after the first round) each decision maker makes a new decision with probability $q = 0.1$, and repeats his last choice with probability $1 - q$.

We use computer simulations to derive the behavior of 20 virtual agents in game 1 of Table 2 over 200 rounds. The agents are matched so that each agent is matched to every other agent once in 20 rounds. We produce 50 artificial data sets in this manner. We then conduct the two sets of analysis described in Section 3

on these artificial data sets with the original EWA model. The results reproduce the pattern observed in Section 3: For the 50 artificial data sets, the median estimated EWA parameters in within-game analysis are $N(0) = 0$, $\delta = 0$, $\phi = 0.10$, $\kappa = 0.01$ and $\lambda = 0.07$. The new-game analysis, on the other hand, produces the median parameter estimates $N(0) = 0$, $\delta = 1$, $\phi = 0.49$, $\kappa = 0.49$, and $\lambda = 0.06$, which are remarkably close to the underlying parameter values.

Note that the within-game analysis implies virtually pure reinforcement learning (which we interpret as high inertia) while new-game analysis implies pure belief learning (or alternatively low inertia). The misspecification of the model (ignoring the fact that $q < 1$), results in biased estimates in the within-game analysis but almost no bias in the new-game estimation. This is because the misspecification has little impact on the aggregate choice frequencies, which are the focus of the new-game estimation.

5. General discussion

The results presented in this work demonstrate the robustness of the observations that led previous descriptive learning studies to reach seemingly contradictory conclusions in various streams of the learning literature. The main contribution of this work is the demonstration that when the differences between new- and within-game predictions are taken into account, the various observations can coexist. The co-existence implies that the results are not inconsistent and that they can be summarized with coherent conclusions.

Two of the observations replicated here concern the value of the assumption of general parameters across situations. The first and foremost observation highlights the usefulness of this assumption. With this assumption the descriptive learning models we considered outperform the equilibrium predictions even when their parameters are estimated based on different sets of games. The second observation highlights a limitation of general parameters assumption: When future behavior is predicted based on past behavior (within-game predictions), the assumption of general parameters over games (and individuals) is counterproductive.

Two additional observations concern the role of action inertia. In within-game predictions, action inertia appears to be one of the most important predictors. Models that allow for action inertia clearly outperform other models in within-game predictions. Under the assertion that in within-game estimation of EWA, reinforcement learning captures action inertia, this observation can explain the importance of adding reinforcement learning to belief learning in EWA (Camerer & Ho, 1999). Wilcox (2005), for example, showed that when individuals are simulated from EWA

⁸Indeed, most studies that focus on within game predictions relax some of the constraints (e.g., Stahl, 1996; Cheung & Friedman, 1997; Bussemeyer & Stout, 2002).

with heterogeneous parameters and the data is then fitted with pooled parameters over individuals, the results unduly favor reinforcement learning. In Wilcox's results, higher heterogeneity in the underlying data results in lower estimated δ . Action inertia also explains the advantage of models of reinforcement sum over reinforcement average in Chen and Tang (1998).

However, when new-game predictions are derived, action inertia does not appear to be an important factor. This result is consistent with Erev et al.'s (1999) observation that the average reinforcement model outperforms the reinforcement sum model in new-game predictions.

The final observation involves sufficient conditions for the current results. As demonstrated in Section 4, the pattern highlighted here can be a product of a relatively minor misspecification of the current models. One sufficient misspecification involves the possibility that human subjects do not make a new decision in each trial. This possibility may seem inconsequential, and indeed it may have little effect on aggregate choice proportions. Nevertheless, it has substantial effect on the parameter estimates in within-game analysis with the models considered here. This effect can impair the new-game predictive power of models estimated in within-game analysis.

5.1. Implications and future research

We believe that the observations summarized above demonstrate the value of two distinct research goals. One goal is improving the models studied here in order to avoid misspecifications that create contradictions between the results of within- and new-game analyses. The current research suggests that a more careful modeling of action inertia can be a constructive step toward this goal. However, additional steps will be necessary. The current research does not address all the contradictions in the learning literature; it merely demonstrates that two of the important contradictions are related to the distinction between within- and new-game predictions.

The second research goal supported by the current analysis involves explicit distinction, at least initially, between within- and new-game predictions. This distinction is particularly appealing under the assertion that it is unlikely that a model accurate enough to eliminate the differences between the two prediction tasks will be found soon. Repetition is only one of many cognitive processes that can dramatically affect sequential dependencies without a large effect on the aggregate choice rate. Other known examples include alternation and negative recency (see Rapoport, Erev, Abraham, & Olson, 1997). An examination of individual data shows that different individuals exhibit different type of sequential dependencies. Thus, it may not be easy to

develop a model of sequential dependencies that is sufficiently parsimonious to derive new-game predictions.

Moreover, the distinction between new- and within-game predictions may not too costly. Similar distinctions are made in other applied sciences. For example, in meteorology different models are used to predict the weather tomorrow and next year (e.g., Vintzileos, Delecluse, & Sadourny, 1999). Interestingly, as in the learning case, action inertia (the weather today) variables are useful in the very near future but not in predicting farther ahead.

The distinction between new- and within-game predictions can be used to facilitate the development of practical applications of current learning research. Specifically, there are many natural situations in which good within-game predictions can be useful, and many natural situations in which good new-game predictions can be useful. One example of a situation in which within-game predictions would be useful involves the task of developing efficient educational software. Software developers can use learning models in order to predict the next response of a student based on past responses. It is easy to see that in this setting, models that assume action inertia and avoid the constraint of general parameters over tasks and agents are likely to be most effective. Another example involves the prediction of future production time (see e.g., MacLeod, 2001; Jovanovic & Nyarko, 1996).

New-game predictions are likely to be useful during the design of a new incentive structure. Examples include the attractiveness of a new slot machine (Haruvy et al., 2001), new rule enforcement strategies (Perry et al., 2002; Erev et al., 2004), and new promotion methods (Haruvy & Erev, 2004).

5.2. Summary

The current analysis can be summarized with two main assertions. The first is pessimistic. It implies that the knowledge accumulated thus far in the experimental learning literature has not led us to a discovery of a general model of learning. It seems that the most appropriate model in a given situation depends on the prediction task and the available information.

The second suggestion is more optimistic. It suggests that at least one inconsistency in the literature might be a result of ignoring the effect of the type of available information on a model's success. When information concerning past behavior in the relevant game is available, this information should be used. The accumulated knowledge implies that information on past behavior in a particular game is best utilized by models that assume high action inertia and allow for game and individual specific parameters. On the other hand, when trying to predict behavior under a new incentive

structure, researchers have to rely on generalizations of the relationship between the incentive structure and behavior. The current results suggest that good generalizations may be provided with models that assume general parameters across situations and low-action inertia.

Acknowledgments

This research was supported by a grant from the USA National Foundation of Science and the USA–Israel Binational Science Foundation. We thank Al Roth, Greg Barron, Bob Slonim, Amnon Rapoport, Dale Stahl, Drew Fudenberg, Colin Camerer, Teck-Hua Ho, John Van Huyck, Barry Sopher, Yaw Nyarko, Andrew Schotter, Charles Holt, Barry O’Neill and the participants of seminars at Harvard, Columbia, NYU, University of Michigan, University of Texas at Austin, University of North Carolina, Universitat Pompeu Fabra, University of Chicago, Florida State University, the Economic Science Association meetings, and the Florence workshop on learning and bounded rationality in markets. We are also indebted to three anonymous referees for their help in improving this paper.

References

- Amemiya, T. (1985). *Advanced econometrics*. Cambridge: Harvard University Press.
- Anderson, S., Goeree, J., & Holt, C. (2001). Minimum-effort coordination games: Stochastic potential and logit equilibrium. *Games & Economic Behavior*, 34(2), 177–199.
- Bereby-Meyer, Y., & Erev, I. (1998). On learning to become a successful loser: A comparison of alternative abstractions of learning processes in the loss domain. *Journal of Mathematical Psychology*, 42, 266–286.
- Blume, A., DeJong, D. V., Lowen, A., Neumann, G. R., & Savin, N. E. (2002). *Learning and experiments: The bootstrap to the rescue*. Working paper, University of Pittsburgh.
- Bolton, G. (1991). A comparative model of bargaining: Theory & evidence. *American Economic Review*, 81(5), 1096–1136.
- Brown, J. N., & Rosenthal, R. W. (1990). Testing the minimax hypothesis: A re-examination of O’Neill’s game experiment. *Econometrica*, 58, 1065–1081.
- Busemeyer, J. R. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed sample, and sequential sampling models. *Journal of Experimental Psychology*, 11, 538–564.
- Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment*, 14(3), 253–262.
- Busemeyer, J. R., & Wang, Y. M. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, 44(1), 171–189.
- Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.
- Camerer, C., & Ho, T.-H. (1998). Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity, and time-variation. *Journal of Mathematical Psychology*, 42, 305–326.
- Camerer, C., & Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4), 827–874.
- Capra, M., Goeree, J. K., Gomez, R., & Holt, C. (2002). Learning and noisy equilibrium behavior in an experimental study of imperfect price competition. *International Economic Review*, 43, 613–636.
- Chen, Y., & Gazzale, R. (2004). When does learning in games generate convergence to Nash equilibria? The role of supermodularity in an experimental setting. *American Economic Review*, 94, 1505–1535.
- Chen, Y., & Khoroshilov, Y. (2003). Learning under limited information. *Games and Economic Behavior*, 44, 1–25.
- Chen, Y., & Tang, F.-F. (1998). Learning and incentive compatible mechanisms for public goods provision: An experimental study. *Journal of Political Economy*, 106, 633–662.
- Cheung, Y.-W., & Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, 19, 46–76.
- Cooper, D. J., & Kagel, J. H. (2002). *Learning and transfer in signaling games*. Mimeo, Case Western Reserve University.
- Crawford, V. (1985). Learning behavior and mixed strategy Nash equilibria. *Journal of Economic Behavior and Organization*, 6, 69–78.
- Daniel, T. E., Seale, D. A., & Rapoport, A. (1998). Strategic play and adaptive learning in sealed bid bargaining mechanism. *Journal of Mathematical Psychology*, 42, 133–166.
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, forthcoming.
- Erev, I., Bereby-Meyer, Y., & Roth, A. (1999). The effect of adding a constant to all payoffs: Experimental investigation, and implications for reinforcement learning models. *Journal of Economic Behavior and Organization*, 39(1), 111–128.
- Erev, I., Ingram, P., Raz, O., & Shany, D. (2004). On the possibility of gentle rule enforcement: The example of cheating in exams. Mimeo, Technion.
- Erev, I., & Roth, A. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique mixed strategy equilibria. *The American Economic Review*, 88(4), 848–881.
- Erev, I., Roth, A., Slonim, R., & Barron, G. (2002). Predictive value & the usefulness of game theoretic models. *International Journal of Forecasting*, 18, 359–368.
- Feltovich, N. (2000). Reinforcement-based vs. beliefs-based learning in experimental asymmetric-information games. *Econometrica*, 68(3), 605–641.
- Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games*. Cambridge, MA: MIT Press.
- Haruvy, E. (1999). Initial conditions, adaptive dynamics, and final outcomes. Doctoral dissertation, University of Texas.
- Haruvy, E., & Erev, I. (2004). *Price promotions and reinforcement learning*. Mimeo.
- Haruvy, E., Erev, I., & Sonsino, D. (2001). The medium prizes paradox: Evidence from a simulated casino. *Journal of Risk and Uncertainty*, 22(3), 251–261.
- Haruvy, E., & Stahl, D. O. (1999). *Initial conditions, adaptive dynamics, and final outcomes*. Working paper, University of Texas.
- Haruvy, E., & Stahl, D. O. (2004). Deductive versus inductive equilibrium selection: Experimental results. *Journal of Economic Behavior and Organization*, 53(3), 319–331.
- Hopkins, E. (2002). Two competing models of how people learn in games. *Econometrica*, 70, 2141–2166.
- Jovanovic, B., & Nyarko, Y. (1996). Learning by doing and the choice of technology. *Econometrica*, 64, 1299–1310.
- Kitzis, S., Kelley, H., Berg, E., Massaro, D. W., & Friedman, D. (1998). Broadening the tests of learning models. *Journal of Mathematical Psychology*, 42, 327–355.
- MacLeod, B. (2001). Uncertainty, planning costs & the theory of learning by doing. USC working paper.

- McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1), 6–38.
- Mookherjee, D. M., & Sopher, B. (1997). Learning and decision costs in experimental constant sum games. *Games and Economic Behavior*, 19(1), 97–132.
- Nagel, R., & Tang, F. F. (1998). Experimental results on the centipede game in normal form: An investigation on learning. *Journal of Mathematical Psychology*, 42, 356–384.
- Nyarko, Y., & Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica*, 70, 971–1006.
- O'Neill, B. (1987). Nonparametric test of the minimax theory of two-person zerosum games. *Proceedings of the national academy of sciences*, 84(7), 2106–2109.
- Perry, O., Haruvy, E., & Erev, I. (2002). Frequent delayed probabilistic punishment in law enforcement. *Economics of Governance*, 3(1), 71–86.
- Rapoport, A., & Boebel, R. B. (1992). Mixed strategies in strictly competitive games: A further test of the minmax hypothesis. *Games and Economic Behavior*, 4, 261–283.
- Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma*. Ann Arbor: University of Michigan Press.
- Rapoport, A., Daniel, T. E., & Seale, D. A. (1998). Reinforcement-based adaptive learning in asymmetric two-person bargaining with incomplete information. *Journal of Experimental Economics*, 1(3), 221–253.
- Rapoport, A., Erev, I., Abraham, E. V., & Olson, D. E. (1997). Randomization and adaptive learning in a simplified poker game. *Organizational behavior and Human Decision Processes*, 69, 31–49.
- Roth, A., & Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games & Economic Behavior*, 8(1), 164–212.
- Salmon, T. (2001). An evaluation of econometric models of adaptive learning. *Econometrica*, 69(6), 1597–1628.
- Samuelson, W., & Zeckhauser, R. (1988). Status quo in decision-making. *Journal of Risk and Uncertainty*, 1, 7–59.
- Shachat, J. (2001). Mixed strategy play and the minimax hypothesis. *Journal of Economic Theory*, 104, 189–226.
- Stahl, D. (1996). Boundedly rational rule learning in a guessing game. *Games and Economic Behavior*, 16(2), 303–330.
- Van Huyck, J., Cook, J., & Battalio, R. (1997). Adaptive behavior and coordination failure. *Journal of Economic Behavior and Organization*, 32(4), 483–503.
- Vintzileos, A., Delecluse, P., & Sadourny, R. (1999). On the mechanisms in a tropical ocean-global atmosphere coupled general circulation model. Part I: Mean state and the seasonal cycle. *Climate Dynamics*, 15(1), 43–62.
- Wilcox, N. (2005). *Heterogeneity and learning principles*. Working paper.

Further reading

- Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics*, 112, 631–645.
- Haruvy, E., & Erev, I. (2001). Interpreting parameters in learning models. In R. Zwick, & A. Rapoport (Eds.), *Experimental business research*. Dordrecht: Kluwer Academic Publishers.
- Ho, T.-H., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best-response in experimental P-beauty contests. *The American Economic Review*, 88(4), 947–969.
- Myers, J. L., Suydam, M. M., & Gambino, B. (1965). Contingent gains & losses in risk-taking situations. *Journal of Mathematical Psychology*, 2, 363–370.
- Rosenthal, R. W., Shachat, J., & Walker, M. (2003). Hide and seek in Arizona. *International Journal of Game Theory*, 32(2), 273–293.
- Sarin, R., & Vahid, F. (2001). Predicting how people play games: A simple dynamic model of choice. *Games and Economic Behavior*, 34(1), 104–122.
- Selten, R., & Buchta, J. (1998). Experimental sealed bid first price auctions with directly observed bid functions. In D. Budescu, I. Erev, R. Zwick (Eds.), *Games and human behavior: Essays in honor of Amnon Rapoport* (pp. 53–78).