

**Ido Erev · Alvin E. Roth · Robert L. Slonim ·
Greg Barron**

Learning and equilibrium as useful approximations: Accuracy of prediction on randomly selected constant sum games

Received: 1 September 2005 / Revised: 10 January 2007 /
Published online: 24 February 2007
© Springer-Verlag 2007

Abstract There is a good deal of miscommunication among experimenters and theorists about how to evaluate a theory that can be rejected by sufficient data, but may nevertheless be a useful approximation. A standard experimental design reports whether a general theory can be rejected on an informative test case. This paper, in contrast, reports an experiment designed to meaningfully pose the question: “how good an approximation does a theory provide on average.” It focuses on a class of randomly selected games, and estimates how many pairs of experimental subjects would have to be observed playing a previously unexamined game before the mean of the experimental observations would provide a better prediction than the theory about the behavior of a new pair of subjects playing this game. We call this quantity the model’s equivalent number of observations, and explore its properties.

JEL Classification Numbers C72 · C90

This research was supported by a grant from the U.S. National Science Foundation and the USA–Israel Binational Science Foundation. We are very grateful for helpful conversations with David Budescu, Jerry Busemeyer, Gary Chamberlain, Paul Feigin, Dave Krantz, Jack Porter, Tom Wallsten, Wolfgang Viechtbauer, and Richard Zeckhauser.

I. Erev

Faculty of Industrial Engineering and Management, Technion,
Minerva Center for Cognitive Studies, Haifa 32000, Israel
E-mail: erev@tx.technion.ac.il

A. E. Roth (✉)

Department of Economics, Harvard University, Cambridge, MA, USA
E-mail: aroth@hbs.edu

R. L. Slonim

Department of Economics, Case Western Reserve University, Cleveland, OH, USA

A. E. Roth · G. Barron

Harvard Business School, Boston, MA, USA

1 Introduction

Conventional experimental methods are well suited to testing whether theories are false—i.e., if their predictions can be rejected. A single well-chosen example, carefully examined, is sufficient to demonstrate, for instance, that not every game quickly elicits equilibrium behavior. Such a design is less well suited to evaluating whether a theory, although false in this sense, might nevertheless provide a useful approximation for predicting behavior on average. Sometimes this has caused miscommunication between economists and experimental psychologists, and even between experimental economists and theorists.¹

The study of the descriptive power of equilibrium and learning models is one of the longest running and most important examples of this miscommunication. Early experimental research by psychologists on equilibrium of zero-sum and other simple matrix games revealed that equilibrium could be rejected. Estes (1950), Bush and Mosteller (1955), Luce (1959), Suppes and Atkinson (1960) and others proposed as alternatives simple learning models that could not be rejected by these early experiments. However, larger and more varied data sets eventually led to the rejection of these learning models also (see e.g., Siegel et al. 1964; Edwards 1961), and interest waned.

O'Neill (1987) critiqued earlier experimenters' conflicting evidence concerning minimax play, and presented an experiment involving a game in which behavior seemed to conform closely to the mixed-strategy equilibrium prediction. Brown and Rosenthal (1990) reanalyzed O'Neill's data and showed the minimax hypothesis could be rejected, and O'Neill (1991) replied that showing that a theory is not "exactly correct" is different from showing that it is not close. It is exactly this kind of debate that we hope to make more precise.

Interest in learning models has been stirred again among economists by the demonstration that simple learning models could account for behavior in a variety of games (e.g., Roth and Erev 1995), including matrix games having unique equilibrium in mixed strategies (Erev and Roth 1998). Selten and Chmura (2005) show that there are also static models that can predict data from constant sum games better than equilibrium.

One reason that many different investigators have in the past reached different conclusions about equilibrium behavior is that they experiment on or analyze behavior from different games. Erev and Roth (1998) addressed this by examining all the published experiments with at least 100 periods of play of games with a unique mixed strategy equilibrium. They observed that simple learning models could provide more accurate predictions than equilibrium over these data. But there was an experimenter effect in the games used in prior experiments, with equilibrium predicting behavior substantially less well in the games studied by some experimenters than others. Because the previously studied games were not

¹ More than one psychologist has wondered: Are economists simply bad scientists? (see e.g., Tversky 1996, and the reply by Roth 1996). In this connection, for example, the robustly replicable deviations from expected utility maximization observed by Allais (1953) showed early in the history of expected utility theory that not all people acted like expected utility maximizers all the time. But the continued use of expected utility theory, even in the face of proposed alternatives, is presumably related to the sense among many users of the theory that it is a useful approximation, even if not precisely true.

a random sample of games, no conclusions could be drawn from them about how close various predictions might be on average to observed behavior.

The present paper is intended to recast these various discussions by using an experimental design that focuses on the ability of a model to predict behavior well on average in a specified universe of games. For this reason we will look at a *random* sample of two-player two-action constant sum games.² Each pair of subjects will play a single game. Multiple pairs of subjects will play each game for 500 periods so that the sample mean and variance of observed play can be estimated over various time scales. These estimates will be compared with the predictions of (minmax) equilibrium, Selten and Chmura's impulse balance equilibrium, and of various learning models.

Looking at how theoretical predictions compare with observed behavior over a random selection of games will allow us to begin to say something about how well the predictions compare to average behavior on that class of games. Nevertheless, each of the models we look at is false in the sense that the null hypothesis that it is precisely correct can be rejected. So we are still left with the question of how to assess how useful these theoretical models are as approximations (which will of course depend on the use to which we plan to put them). We will address this question by examining how well on average each model performs the task of predicting the proportion of times each action is chosen by each player, for a previously unobserved pair of subjects playing one of the games.³

1.1 Equivalent number of observations (ENO)

We will compare the predictions of each model with another way to predict how a new pair of subjects will play one of the games, which is to look at the average behavior of other pairs of subjects who played the same game. The more subjects who have already played the game, the better the estimate that past play will give of the mean behavior from this subject population on this game. So one way to measure how useful is the prediction of a particular model is to ask how many prior observations of subjects playing the game would be needed to make as accurate a prediction as the model. We will call this measure the model's *equivalent number of observations* (ENO).⁴ The ENO of a model is closely related to how one would combine the prediction of the model with the observed data to obtain a new prediction: a regression would weight the prior observations and the prediction of the model equally when there are $n = \text{ENO}$ prior observations.

We will look at the ENO of equilibrium and the other models we consider from the point of view of predicting the observed choices in all 500 periods of play, and also in the first 100 periods and the last 100 periods, since it might be expected that learning models will do best at predicting the early periods of play while equilibrium will do best at predicting play once the players have gained experience playing with one another. To foreshadow our results (in Table 2), we estimate from

² As far as we are aware, this is the first experiment to study a random sample of games.

³ Thus, we use these models to examine the proportion of times each action is chosen. We could also examine many other observable statistics about which the models make predictions; Slonim et al. (2006) for instance look at other statistics including variance and serial correlation.

⁴ In Erev et al. (2002) we refer to this measure as "predictive value".

the data that equilibrium has an ENO of less than 1 for predicting the first 100 periods, or all 500 periods, and an ENO of between 1 and 2 for predicting the last 100 periods of play. This means that, when the task is to predict the behavior of a previously unobserved pair of subjects playing a game from this universe of games, observing the behavior of a single pair of other players will provide a better prediction than equilibrium for the choices of the new pair of players in their first 100 periods of play, but the equilibrium will provide a better prediction for the last 100 periods. However, the average proportion of choices of *two* pairs of players playing the same game will provide a better prediction than equilibrium even for the last 100 periods of play.

The other models we investigate all have higher estimated ENOs than equilibrium, not only when the players are inexperienced, but also when they are experienced. The models with the highest ENO for predicting the first 100 periods of play are simple reinforcement learning models, with ENOs of just over 37. This means that they provide a better prediction for how a pair of subjects will play in the first 100 periods than even a fairly large sample of other pairs of subjects who played the same game. For the last 100 periods, all the models we consider other than equilibrium have ENO's greater than 10; i.e., they all provide better predictions than could be obtained from observing ten pairs of players playing the game in question. A variant of fictitious play has a higher ENO than reinforcement learning, and so does the (static) impulse balance model of Selten and Chmura.

The paper is organized as follows. Section 2 presents the experimental design and reports how closely the various models predict observed behavior, as measured by the mean squared error of the predictions from the data. Mean squared errors are difficult to interpret by themselves, so Sect. 3 develops the notion of a model's ENO, which we motivate by its relation to regression analysis for combining the predictions of the model and the data. We then compare the ENO of the various models, and conclude.

2 The random sample of games, and the experimental design

Each subject played 500 repetitions of one of 40 two-player, constant sum games (shown in Table 1), against a fixed, anonymous opponent. The numbers in each row represent one of the games by showing the probabilities ($\times 100$) that the players will win a fixed amount v on each trial, for each element of the payoff matrix resulting from the players' actions. For example, if in a given period both players choose action "A," then player 1 will win v with the specified probability p_1 listed in column AA, and player 2 will win v with probability $1 - p_1$. A player who does not win v earns zero for that period. In each of the games played, v was set at \$0.04 and a player's payoff from the game was the sum of his payoffs over the 500 periods of play (plus a fixed showup fee). Each player played only one game (defined by the probabilities in one row of Table 1), against a fixed opponent. All transactions were conducted anonymously via networked computers.

Such a game either has a (weakly) dominant strategy for at least one of the players, or has a unique mixed-strategy equilibrium at which both players play each of their strategies with positive probability. In each random sample described below, the probabilities p_1 through p_4 were independently chosen from the uniform

Table 1 The 30- and 10-game samples and the observed choice proportions

Payoff matrix					Player 1					Player 2				
Game	AA	BB	BA	BB	1	2	3	4	5	1	2	3	4	5
30 game sample					Choice proportions (in 5 blocks of 100 trials)									
1	58	94	98	51	0.56	0.70	0.47	0.55	0.49	0.47	0.67	0.54	0.36	0.33
2	100	24	5	96	0.58	0.61	0.60	0.69	0.57	0.43	0.49	0.41	0.48	0.49
3	82	37	19	84	0.55	0.49	0.55	0.56	0.66	0.66	0.61	0.58	0.59	0.54
4	55	36	26	59	0.46	0.36	0.31	0.35	0.34	0.58	0.41	0.49	0.56	0.51
5	28	85	78	0	0.64	0.45	0.49	0.45	0.46	0.57	0.49	0.46	0.47	0.48
6	7	97	87	38	0.50	0.39	0.37	0.46	0.40	0.54	0.50	0.51	0.60	0.48
7	70	26	41	87	0.61	0.48	0.41	0.08	0.14	0.61	0.46	0.59	0.61	0.47
8	38	87	100	60	0.31	0.30	0.31	0.22	0.19	0.29	0.36	0.44	0.37	0.19
9	91	46	49	71	0.69	0.58	0.52	0.65	0.64	0.42	0.36	0.33	0.34	0.36
10	98	57	15	93	0.82	0.75	0.75	0.75	0.89	0.32	0.15	0.21	0.25	0.17
11	96	49	05	83	0.90	0.67	0.76	0.89	0.78	0.38	0.44	0.51	0.46	0.34
12	55	39	14	64	0.64	0.72	0.65	0.81	0.79	0.74	0.72	0.76	0.60	0.43
13	80	62	67	100	0.52	0.17	0.76	0.19	0.00	0.58	0.66	0.21	0.80	1.00
14	55	16	8	29	0.63	0.36	0.26	0.20	0.19	0.42	0.28	0.16	0.21	0.13
15	95	3	28	49	0.52	0.33	0.16	0.14	0.13	0.21	0.17	0.19	0.24	0.31
16	51	33	05	97	0.98	0.91	0.90	0.90	0.92	0.87	0.66	0.65	0.34	0.37
17	72	44	60	66	0.46	0.30	0.26	0.25	0.29	0.25	0.26	0.16	0.13	0.04
18	48	27	2	36	0.55	0.58	0.44	0.49	0.32	0.51	0.56	0.45	0.75	0.75
19	59	95	78	04	0.92	1.00	0.97	1.00	0.73	0.43	0.45	0.65	0.76	0.77
20	88	15	56	69	0.30	0.22	0.06	0.28	0.33	0.32	0.64	0.58	0.60	0.49
21	57	17	22	31	0.41	0.35	0.36	0.27	0.00	0.19	0.20	0.27	0.21	0.00
22	19	5	1	58	0.58	0.75	0.69	0.75	0.77	0.67	0.60	0.70	0.74	0.71
23	96	27	78	97	0.06	0.06	0.00	0.00	0.00	0.50	0.95	0.89	0.96	1.00
24	100	49	71	76	0.49	0.53	0.65	0.41	0.24	0.34	0.35	0.32	0.33	0.34
25	48	39	0	50	0.93	1.00	0.99	1.00	1.00	0.07	0.00	0.00	0.00	0.00
26	37	97	83	79	0.33	0.06	0.15	0.08	0.21	0.45	0.12	0.22	0.13	0.21
27	35	98	81	79	0.33	0.26	0.13	0.09	0.03	0.60	0.43	0.23	0.10	0.21
28	47	10	46	59	0.40	0.08	0.09	0.00	0.00	0.51	0.39	0.37	0.39	0.20
29	87	0	3	9	0.44	0.28	0.11	0.23	0.20	0.01	0.00	0.00	0.05	0.03
30	25	91	94	90	0.40	0.69	0.36	0.05	0.11	0.63	0.85	0.59	0.29	0.43
10 game sample														
31	77	35	8	48	0.59	0.51	0.63	0.61	0.63	0.31	0.28	0.35	0.34	0.31
32	73	74	87	20	0.80	0.90	0.79	0.85	0.86	0.37	0.44	0.36	0.35	0.28
33	63	8	1	17	0.61	0.57	0.65	0.54	0.54	0.22	0.19	0.25	0.23	0.22
34	55	75	73	60	0.32	0.22	0.26	0.29	0.28	0.62	0.52	0.47	0.47	0.43
35	5	64	93	40	0.35	0.38	0.38	0.38	0.40	0.36	0.36	0.34	0.27	0.27
36	46	54	61	23	0.60	0.61	0.66	0.65	0.67	0.38	0.39	0.42	0.40	0.47
37	89	53	82	92	0.35	0.32	0.27	0.28	0.26	0.45	0.53	0.52	0.54	0.57
38	88	38	40	55	0.52	0.47	0.34	0.34	0.34	0.25	0.19	0.22	0.22	0.24
39	40	76	91	23	0.54	0.59	0.59	0.54	0.55	0.41	0.42	0.51	0.48	0.43
40	69	5	13	33	0.33	0.36	0.35	0.26	0.30	0.20	0.21	0.20	0.20	0.19

Each player was asked to select between A and B. 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 40 games were played by fixed pairs for 500 trials. The right hand columns show the proportion of A choices in five blocks of 100 trial by each of the players

distribution on the values $[0.00, 0.01, \dots, 0.99, 1.00]$. Games generated in this way were included in the sample if they had a unique mixed strategy equilibrium.

Each player knew the probabilities that defined the game he was playing. After each period of play each player learned what action the other player had chosen (and

therefore with what probability each player would receive the payoff $v = \$0.04$). Players also knew whether or not they received the payoff v , but did not know whether the other player received v .

Because the games have binary lottery payoffs, the equilibrium predictions can be determined without estimating any unobservable parameters involving risk aversion (Kagel and Roth 1995; Roth and Malouf 1979; Wooders and Shachat 2001).⁵ Since equilibrium predictions do not require estimating free parameters, a single random sample of games would be adequate for measuring the closeness of the equilibrium prediction to the observed behavior. However, since the learning models have free parameters, which must be estimated, and since we are interested in predictive power for new games, we collect data from two distinct random samples of games; we estimate the free parameters of the learning models from one sample and use these estimates to predict behavior in the other sample.

We examine a 10-game sample and a 30-game sample. Each subject (in either sample) plays one game, against a fixed opponent. Each game in the 30 game sample is played by one pair of subjects. Each game in the ten game sample is played by nine pairs, three each in Boston, Haifa, and Pittsburgh (at the experimental laboratories of Harvard, Technion, and University of Pittsburgh). Because the games are played by multiple subject pairs, the ten-game random sample can be used to assess the ENO of the models' predictions, since each model's predictions can be compared with the predictions from the means of subsamples of player pairs for each game.⁶

2.1 Equilibrium, impulse balance equilibrium, and five learning models

We will consider the predictive value of equilibrium, impulse balance equilibrium (Selten and Chmura 2005) and five learning models of choice behavior. Equilibrium for two-person zero-sum games is one of the oldest ideas in game theory, whose existence was proved by von Neumann (1928). It is a special case of Nash equilibrium (Nash 1950) for general strategic games, in which each player chooses each of his actions with probabilities such that, given the strategy of the other player, no change in probabilities would increase his expected payoff. In zero sum games, a player's equilibrium strategy can be calculated by maximizing the minimum payoff he might get for any action of his opponent. The games in our experiment have only two choices per player, and are randomly chosen from the universe of such games having a unique equilibrium in nontrivial mixed strategies (strategies such that no action is played with certainty).⁷

⁵ And because the games are constant sum, many of the concerns expressed e.g., by Ochs and Roth (1989), Bolton and Ockenfels (2000) or Weibull (2004) about experimental control of other aspects of players' preferences are ameliorated. Furthermore, the repeated game equilibrium can be analyzed in terms of the actions in the stage game, which allows us to avoid the difficulties associated with estimating repeated game strategies (cf. Engle-Warnick and Slonim 2006).

⁶ Unlike the sample of games, the set of subjects is not a random draw from a well specified universe. But the robustness of the results across subjects sampled in this diverse set of cities suggests that these results are not closely dependent on the particular subject pool.

⁷ Mixed strategies are not only the theoretically difficult case that is the focus of von Neumann's minimax theorem, they also constitute a behaviorally difficult test of equilibrium, because at equilibrium no player has a positive incentive to play the equilibrium mixed strategy.

Four of the learning models considered here can be captured with two basic assumptions. The first assumption is a stochastic choice rule in which the probability of selecting action k at trial t is given by

$$P_k(t) = \frac{e^{q_k(t)D(t)}}{\sum_{j=1}^2 e^{q_j(t)D(t)}},$$

where $q_j(t)$ is the propensity to select action j and $D(t)$ modulates how decisive is the decision maker. (When $D(t)$ equals 0, each action is chosen with equal probability; when it is large, the action with higher propensity is chosen with high probability.)

The second assumption concerns the adjustment of propensities as experience is gained. The propensity to select action k at trial $t + 1$ is a weighted average of $q_k(t)$, the propensity at t , and $v_k(t)$, the payoff from selecting this strategy at t

$$q_k(t + 1) = [1 - W(t)] \cdot q_k(t) + W(t)v_k(t).$$

The initial value, $q_k(1)$ is assumed to equal $A(1)$ —the expected payoff from random choice [e.g., $A(1)$ of Player I in Game 1 is $(58 + 94 + 98 + 51)/4 = 75.25$]. The models differ with respect to the decisiveness function $D(t)$, the weighting function $W(t)$ and the assumed value of the obtained payoff $v_k(t)$.

The first model, referred to as reinforcement learning (RL) assumes stable payoff sensitivity, $D(t) = \lambda$, and insensitivity to forgone payoff: $W(t) = w$ (i.e., a constant) if k was selected at t , and 0 otherwise. In addition, this model assumes that $v_k(t)$ is the realized payoff (0 or 100 points, depending on the outcome of the binary lottery).

The second model, referred to as normalized reinforcement learning (NRL), is similar to the model proposed by Erev et al. (1999). It is identical to RL with one exception: payoff sensitivity is assumed to decrease with payoff variability. Specifically,

$$D(t) = \frac{\lambda}{S(t)},$$

where $S(t)$ is the weighted average of the difference between the obtained payoff at trial t and the maximal recent payoff:

$$S(t + 1) = (1 - w)S(t) + w|\max(\text{recent}_1, \text{recent}_2) - v_k(t)|,$$

where recent_i is the most recent observed payoff from action i (it equals $v_k(t)$ if $i = k$). Before the first observation of the payoff of action i , $\text{recent}_i = A(1)$. The initial value, $S(1)$, is assumed to equal λ .

The third model is stochastic fictitious play (SFP, see Fudenberg and Levine 1998; Goeree and Holt 1999; Cheung and Friedman 1997; Cooper et al. 1997). It assumes stable payoff sensitivity, $D(t) = \lambda$, and sensitivity to forgone payoff $W(t) = w$. To capture the logic behind the fictitious play rule this model assumes that $v_k(t)$ is the expected payoff in the selected cell (e.g., when both players select A in Game 1, $v_k(t)$ of Player I is 58, independent of the outcome of the binary lottery).

For this reason, the earlier experiments referred to above also concentrated on games with mixed strategy equilibria.

The fourth model, referred to as normalized fictitious play (NFP), was proposed in Ert and Erev (2007) to capture choice behavior in individual decision tasks. It is identical to SFP with the exception of the payoff sensitivity function. Like NRL it assumes $D(t) = \lambda/S(t)$.

The fifth model considered here is a non-linear combination of reinforcement learning and fictitious play introduced by Camerer and Ho (1999). This model, referred to as experience weighted attraction (EWA), is described in the Appendix.

Besides the five learning models and equilibrium, the final model we consider, called impulse balance equilibrium, is a static model proposed by Selten and Chmura (2005) to capture behavior of experienced players. We examine this model because it has been shown to provide a surprisingly good approximation of aggregate behavior in a wide set of games without fitting parameters to individual games. This model assumes that observed choice proportions reflect a balance between the “subjective incentives.” The original version of the model does not have free parameters. But the authors assume higher sensitivity to losses (relative to a reference point) than to gains. This assumption is abstracted here with a free parameter, φ , that captures the tendency to overweight losses. The basic model is presented in Appendix 1. Under our formulation, the original version of the model assumes $\varphi = 2$.

Under the assumption of uniform initial propensities, $q_k(1) = A(1)$, the first four learning models have two free parameters, EWA has five parameters, and impulse balance has one parameter.

2.2 Model comparison

To compare the equilibrium predictions with the predictions of the other models we used the generalization criterion (see Busemeyer and Wang 2000): the behavior in the 30 game sample is used to estimate the models’ parameters, and the behavior in the 10 game sample is used to compare the predictions. The analysis focuses on the predictions of aggregate choice probability over all 500 periods of each player in each game, and over the first 100 and last one hundred periods. The right hand columns of Table 1 present these proportions for the 40 games.

A simulation-based grid search with a mean squared error (MSE) criterion was used to estimate the parameters. The predictions of each of the models were derived for a large number of parameter sets. For each parameter set and model we simulated 100 plays of each game and then determined the mean proportion of plays of each action. The simulations were run using the SAS[®] software. The program (and the necessary data) used to estimate the predictions of the NFP model can be found in Erev (2006). The parameter set that minimized the MSE between this prediction and the observed choices over all 500 periods of the 30 game sample were selected. (For two of the models, NFP and impulse balance equilibrium, we also present the predictions with the parameters used in previous studies of these models).

Table 2 shows the MSE’s of the different models. The MSE’s are presented both for the 30 game sample (on which the models’ parameters were estimated) and the 10 game sample. So the 30 game sample shows how closely the data can be fitted by each of the learning models, and the 10 game data shows how closely their predictions match the observed behavior. Recall that MSE’s are calculated as

Table 2 The MSE and ENO scores of the different models

Model	MSE and estimated parameters in the 30 game sample		MSE in the 10-game sample				ENO	
	All 500 periods	All 500 periods	All 500 periods	First 100 periods	Last 100 periods	All 500 periods	First 100 periods	Last 100 periods
Reinforcement learning (RL)	0.0223	0.0274	0.0274	0.0329	0.0414	13.6	37.4	12.67
Normalized reinforcement learning (NRL)	$\lambda = 8, w = 0.25$ 0.02150 $\lambda = 5, w = 0.6$	0.0273	0.0273	0.0329	0.0399	14.0	37.8	25.15
Stochastic fictitious play (SFP)	0.02391 $\lambda = 12, w = 0.05$	0.0287	0.0287	0.0353	0.0407	8.9	9.77	16.3
Normalized fictitious play (NFP)	0.02701 $\lambda = 2.5, w = 0.25$	0.0274	0.0274	0.0365	0.0387	13.6	7.10	107.1
NFP with Ert and Erev parameters	0.03031 $\lambda = 2.75, w = 0.52$	0.0281	0.0281	0.0346	0.0414	9.92	12.35	12.5
Experience weighted adjustment (EWA)	0.02110 $N(1) = 1, \lambda = 6,$ $\delta = 0.3, \rho = 0.16,$ $\phi = 0.2$	0.0284	0.0284	0.0366	0.0408	8.7	6.95	15.4
EWA with the constraint $\phi = 1$	0.02332 $N(1) = 0.001, \lambda = 10,$ $\delta = 0.9, \rho = 1, (\phi = 1)$	0.0301	0.0301	0.0406	0.0411	5.54	3.75	14.02
Impulse balance equilibrium (IBE)	0.0333 $\phi = 1.5$	0.0293	0.0293	0.0388	0.0395	6.78	4.74	31.1
IBE with Selten and Chmura's gain sensitivity parameter	0.0338 $\phi = 2$	0.0303	0.0303	0.0415	0.0397	5.31	3.38	28.9
Equilibrium	0.0389	0.0572	0.0572	0.0750	0.0666	0.79	0.74	1.36

The second column presents the MSE score of the fitted models on the 30-game sample. The parameters were estimated (with the two exceptions noted in the table) to fit the aggregated choice proportion over the 500 periods. The central columns show the MSE of the predictions of the models to the 10-game sample. The analysis focuses on three statistics: choice proportion over the 500 trials, in the first 100 trials, and in the last 100 trials. The right hand columns show the pooled variance (S^2) and the implied ENO scores

the average of the squared difference between each model's prediction and each subject's average choice, so that smaller MSE's indicate better fit.

The predictions of all of the models considered are closer to the observed behavior than is the equilibrium prediction, on each of the time scales considered.

Comparison of the various learning (and impulse balance) models suggests that the difference between the models is not large. The MSE scores for the 30 game sample (column 2 of Table 2) show that the model with the largest number of parameters (EWA) can be most closely fitted to the 30 game data.⁸ But examination of the MSE scores for the ten game sample (columns 3–5 of Table 2) shows the difference between fitting data ex post and predicting new data: EWA predicts slightly less well than the similar models with fewer parameters. Overall, the MSE scores suggest that the difference between the learning models is much smaller than the difference between the models and equilibrium.⁹

3 The usefulness of the approximations

Until now our analysis has focused on the raw distances (MSEs) of the predictions from the data, and the comparison of models this allows.¹⁰ But since all of these models can be rejected (a la Brown and Rosenthal), this still leaves us pretty much where their debate with O'Neill left off: some models are closer to the data than equilibrium, but perhaps equilibrium is "close enough." In what follows we offer a way to more clearly interpret the magnitudes involved.

Specifically, we highlight two related interpretations of a transformation of MSE scores. The first interpretation is the "equivalent number observations" (ENO) of the relevant model. To define this concept formally, we assume that our goal is to predict the behavior of a particular person in a particular experimental condition (e.g., the proportion of *A* choices by Player I in a game drawn from the population of games considered above). Assume further that we have available two forecasts: (1) the prediction of a particular model (one of the models considered above), and (2) the mean behavior of *m* other individuals in the same condition (*m* other player I's in the same game). The ENO of the model is the value of *m* (the size of the experiment) that is expected to lead to a prediction that is as accurate as the model's prediction.

The second interpretation involves the task of combining the two forecasts considered above. We show that under reasonable assumptions the ENO statistic provides an approximation of the optimal weighting of the two forecasts.

⁸ To simplify the presentation and analysis we estimated the parameters over all 500 periods. Thus, the most interesting analysis involves this statistic.

⁹ This observation is not a result of careful selection of the learning models. Similar findings were obtained with many other learning models including the reinforcement learning models studied in Erev and Roth (1998) and Erev et al. (1999), the payoff assessment model (Sarin and Vahid 2001) and a more complex model that assumes reinforcement learning among cognitive strategies (RELACS, Erev and Barron 2005). The similarity of the learning models is consistent with the observations made in Hopkins (2002) and Salmon (2001). Erev and Barron (2005) note that the differences among the learning models becomes much more apparent when the data sets include tasks with limited feedback and significant low probability outcomes.

¹⁰ Erev and Haruvy (2005) discuss the differences between MSE's and maximum likelihood estimators for comparing approximations.

3.1 A simplified problem and derivation

To clarify the derivation of the ENO statistic it is convenient to start with an analysis of a simplified problem. Consider a situation in which an observer (a potential user of the model) can observe the mean behavior of m individuals and is then asked to predict the behavior of individual $m + 1$. The model is assumed to be general; it is used to address a wide class of possible experimental conditions. Each experimental condition in this setting is associated with a true mean μ_i , and a normal distribution (with variance σ_ε^2) around this mean. That is, the results of running condition i can be described as sampling from the normal distribution $N(\mu_i, \sigma_\varepsilon^2)$. The model is assumed to provide approximations that may nevertheless be rejected as the data generating process given sufficiently many observations. Accuracy can thus be enhanced by a careful estimation of the parameters of the model. Using these estimates, the prediction of the mean of experimental condition i by the general model is $G_i = \mu_i + \alpha_i$ where α_i is an error term (the bias of the model in experimental condition i).

As noted by Granger and Ramanathan (1984) the leading methods of combining forecasts of normally distributed variables can be described as variants of regression analyses. To place our discussion of a model's ENO in context, we first discuss three variants of regression that differ with regard to the assumptions made during the estimations of the relevant weights.

3.1.1 Regression, recycling and restrictions

Under traditional regression analysis, the best (least squared difference) prediction of observation $m + 1$ in experimental condition i ($x_{i,m+1}$) based on the two forecasts (G_i and $\bar{X}_{i_first_m}$) is provided by estimating the three free parameters (β_0 , β_1 , and β_2) in the following equation:

$$x_{i,m+1} = \beta_0 + \beta_1(G_i) + \beta_2(\bar{X}_{i_first_m}) + \varepsilon_{i,m+1}. \quad (1)$$

To estimate these parameters we simply need to run multiple experiments in which we can observe the relationship between the criterion variable $x_{i,m+1}$ and the two forecasts. To facilitate robust estimation, these experiments should involve randomly selected experimental conditions from the relevant population of conditions. Under the common assumptions made in regression analysis, Eq. (1) is expected to provide the weighting of the two estimates that minimizes mean squared error.

In the current context, the regression analysis described above has two shortcomings. The first involves the fact that it uses each experimental condition once. One of the $n = m + 1$ observations is used as the dependent variable, and all the other observations are used to compute one of the independent variables, $\bar{X}_{i_first_m}$. Thus, it ignores the symmetry among the n observation. This shortcoming can be addressed with a "regression with recycling" procedure. Under this procedure each of the n observations in each of the N experimental conditions is used as the dependent variable once, and is then "recycled" m times as one of the m other observations used to compute the value of the mean of the other observations. This analysis can use $(n)(N)$ distinct data lines, and the following equation:

$$x_{ij} = \beta_0 + \beta_1(G_i) + \beta_2(\bar{X}_{oij}) + \varepsilon_{ij}, \quad (2)$$

where x_{ij} is observation j in experimental condition i , and \bar{X}_{oij} is the mean of the other m observations (all observations but j) in experimental condition i . Table 4 presents a small data set ($N = 10, m = 3$) of this type, to illustrate the computations.¹¹

A second shortcoming of the traditional regression analysis involves the number of free parameters. It requires the estimation of three different weighting parameters for each value of m . This requirement is likely to be counterproductive when the general model is relatively accurate (i.e., when α_i is small). That is, the estimation of three parameters might lead to over-fitting the data. This risk can be reduced with a variant of “regression with recycling” that uses the restrictions $\beta_0 = 0$, and $\beta_1 + \beta_2 = 1$. We refer to this procedure as “restricted regression with recycling.”

3.1.2 The minimum variance rule and ENO based weighting

As implied by Granger and Ramanathan (1984; and see Gupta and Wilton 1987) the predictions of the restricted regression procedure are identical to the predictions of the minimum variance rule. Under the minimum variance rule (and restricted regression analysis) the optimal weight for the point prediction rule is

$$\hat{W} = \hat{\beta}_1 = \frac{\text{MSE}(\bar{X}_o) - CD(\bar{X}_o, G)}{\text{MSE}(\bar{X}_o) + \text{MSE}(G) - 2CD(\bar{X}_o, G)}, \quad (3)$$

where $\text{MSE}(G)$ is an unbiased estimator of the mean squared error (MSE) of the general model, $\text{MSE}(\bar{X}_o)$ is an unbiased estimator of the MSE of the second predictor (the mean of the other m observations), and $CD(\bar{X}_o, G) = r(\bar{X}_o - x_{ij}, G - x_{ij})\sqrt{\text{MSE}(\bar{X}_o)\text{MSE}(G)}$ is the common deviation where $r(\bar{X}_o - x_{ij}, G - x_{ij})$ is the correlation between the deviations. To clarify the meaning of the relevant terms in this setting, Table 4 shows how they are computed in this 30-observation example.

Under the minimum variance rule the two estimates receive equal weight when $\text{MSE}(\bar{X}_o) = \text{MSE}(G)$. The number of observations used to derive the experiment-based predictions ($m = n - 1$) decreases the error of this prediction ($\text{MSE}(\bar{X}_o)$), but does not have a systematic effect on the error of the model-based prediction ($\text{MSE}(G)$). Thus, it is possible to compute the ENO of the model as the estimate of the size of the experiment for which the two predictors are equally useful and the optimal weight is 0.5. Appendix 2a shows that the exact value is given by

$$\text{ENO} = \frac{S^2}{(M - S^2)}, \quad (4)$$

where S^2 is the pooled variance (over tasks in the experiment), and $M = \text{MSE}(G)$.

¹¹ The data in Table 4 are generated as follows. Each of the ten condition means (μ_i values) were drawn independently from $N(100, 20^2)$. Each of the three observations in each condition (x_{ij} values) were drawn independently from $N(\mu_i, 10^2)$. The prediction of the model for condition i (G_i) was drawn from $N(\mu_i, 5^2)$.

Table 3 Example of a data set that allows regression with recycling

Condition	Depend. variable		Predictors		Deviations		Squared deviations	
	Obs.	x_{ij}	X_{oij}	G_i	$(\bar{X}_{oij} - x_{ij})$	$(G_i - x_{ij})$	$(\bar{X}_{oij} - x_{ij})^2$	$(G_i - x_{ij})^2$
1	1	100	116	104	16	4	256	16
	2	110	111	104	1	-6	1	36
	3	122	105	104	-17	-18	289	324
2	1	124	109	92	-15	-32	225	1024
	2	106	118	92	12	-14	144	196
	3	112	115	92	3	-20	9	400
3	1	102	114	108	12	6	144	36
	2	110	110	108	0	-2	0	4
	3	118	106	108	-12	-10	144	100
4	1	126	117	124	-9	-2	81	4
	2	106	127	124	21	18	441	324
	3	128	116	124	-12	-4	144	16
5	1	62	77	76	15	14	225	196
	2	74	71	76	-3	2	9	4
	3	80	68	76	-12	-4	144	16
6	1	130	136	128	6	-2	36	4
	2	134	134	128	0	-6	0	36
	3	138	132	128	-6	-10	36	100

Table 4 (Contd.)

<i>i</i>	Condition	Obs.	<i>j</i>	<i>x_{ij}</i>	Depend. variable	Predictors		Deviations		Squared deviations		
						\bar{X}_{oij}	G_i	$(\bar{X}_{oij} - x_{ij})$	$(G_i - x_{ij})$	$(\bar{X}_{oij} - x_{ij})^2$	$(G_i - x_{ij})^2$	
7			1	104		100	90	-4	-14	16	196	
			2	112		96	90	-16	-22	256	484	
			3	88		108	90	20	2	400	4	
8			1	70		94	79	24	9	576	81	
			2	90		84	79	-6	-11	36	121	
			3	98		80	79	-18	-19	324	361	
9			1	92		99	94	7	2	49	4	
			2	90		100	94	10	4	100	16	
			3	108		91	94	-17	-14	289	196	
10			1	84		86	76	2	-8	4	64	
			2	82		87	76	5	-6	25	36	
			3	90		83	76	-7	-14	49	196	
										Sum = 4,452	Sum = 4,595	
												$r(\bar{X}_o - x_{ij}, G_i - x_{ij}) = 0.746$

$$MSE(\bar{X}_o) = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{j=1}^n (x_{ij} - \bar{X}_{oij})^2 = \frac{1}{10} \frac{1}{3} 4452 = 148.4$$

$$MSE(G) = M = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{j=1}^n (x_{ij} - G_i)^2 = \frac{1}{10} \frac{1}{3} 4595 = 153.17$$

$$s^2 = MSE(\bar{X}_o) \left(\frac{n-1}{n} \right) = 148.4 \left(\frac{2}{3} \right) = 98.93, ENO = s^2 / (M - s^2) = 98.93 / (153.17 - 98.93) = 1.82$$

$$CD(\bar{X}_o, G) = 0.746 \sqrt{(148.4)(153.17)} = 112.5$$

Variable x_{ij} summarizes the value of observation j in condition i , \bar{X}_{oij} is the mean of the other (two) observations, G_i is the prediction of the general model. The right hand columns and the lower rows present the computation of the statistics discussed in the text

Note that the ENO is a property not only of the model alone but of both the model and the data. On a universe of tasks over which subjects exhibit little variance in behavior, every observation is very informative, so even a very good model (in the sense of being absolutely close to predicting subject's mean behavior) will have a low ENO. In other words, ENO will be larger the more accurately the model predicts subjects' mean behavior, and will also be larger the greater the variation in subjects' behavior.

Under the reasonable assumption that $CD(\bar{X}_o, G) = S^2$ (i.e., that the errors of the two predictors are not correlated), the ENO statistic can be used to provide a simplified approximation of the optimal weighting.¹² The implication of the minimum variance rule in the current setting (after observing m subjects) is (see Appendix 2b):

$$\hat{W} = \hat{\beta}_1 = \frac{\text{ENO}}{\text{ENO} + m}. \quad (5)$$

Equation (5) makes clear why we refer to the "equivalent number of observations" of a theoretical prediction, since, if $\text{ENO} = k$, we give the theoretical prediction the same weight as we would a data set of k observations, when combining it with a data set with m observations.¹³ We refer to the procedure that relies on this observation as "ENO based weighting."

Although the main goal of the procedures proposed here is to consider general models that predict behavior in a universe of conditions, it is easy to see that the ENO statistic can be written as a transformation of conventional statistics used in the study of a single condition. Most importantly, the ENO statistic based on one condition ($N = 1$) and a sample of n observations is closely tied to the t test for the hypothesis that the condition's mean equals 0 (the model $G = 0$) and to Cohen's (1994) *effect size* (d). These relationships are presented in Appendix 2c.

In Erev et al. (2005) we present an extensive comparison of the four combination rules presented above. The analysis shows that when the size of the experiment is small (like the current experiment), the ENO based weighting provides the best prediction.

3.2 The ENO of equilibrium, of impulse balance equilibrium, and of the learning models

In the second row of the three right hand columns of Table 2 we show, for the ten game sample, the average sample variance S^2 . The ENO's of each of the models on

¹² To see why $CD(\bar{X}_o, G) = S^2$ implies independent error note that $CD(\bar{X}_o, G)$ can be written as the sum of the mean squared error of x_{ij} from the relevant population's mean (this error equals S^2), and the common error of the two predictors in predicting the population's mean. Notice that in Table 1's small data set example the estimated value of $CD(\bar{X}_o, G)$ is larger than S^2 (112.5 vs. 98.9). In the data sets we considered the deviation between the estimates decreases when more observations are used in the estimation.

¹³ When the samples are small it is possible to obtain estimates such that $M - S^2 \leq 0$. In restricted regression analysis, these cases lead to the estimation of $\beta_1 > 1$ that implies negative weight to the mean of the first m observations. Since negative correlation between observation $m + 1$ and the first m is impossible in the current setting, we treat these cases as situations in which the ENO of the model is too high to be estimated based on the current sample and set $\hat{W} = 1$.

the ten game sample can then be computed from this figure and the model's MSE, as discussed above, $ENO = S^2/(M - S^2)$. For example, the ENO of the RL model for the first 100 periods of play on the ten game sample is $0.03204/(0.03290 - 0.03204) = 37$. A SAS program that demonstrates the computation of the ENO of the NFP can be found in Erev (2006).

The fact that the ENOs of the models are finite implies that these models only give approximations to the true population means. But the size of the estimated ENOs clearly demonstrates that these models can be useful. For example, the reinforcement and fictitious play models are more useful than our experiment (which observes nine pairs in each game) for predicting the proportion of time additional pairs of players in each game will play each action.

Using ENO as our comparative metric lets us look more closely at the relative performance of equilibrium and the learning models as predictors of the observed behavior. Equilibrium becomes a better predictor as players gain experience with the game. For the first 100 periods, the equilibrium prediction has an ENO less than 1. So, for the first 100 periods, the equilibrium prediction is less informative than observing one pair of players. However by the last 100 periods, equilibrium has an ENO greater than 1. That is, the equilibrium prediction is more informative than observing a single pair of players actually playing the game, if one wishes to predict the behavior of other pairs of players in periods 401 through 500. The trend suggests that the predictive value of equilibrium may increase as players get yet more experience. This may not be too bad for a model with no parameters, but it is sobering to note how much experience is required before equilibrium predicts better than a single observation.

All the learning models and impulse balance equilibrium predict better than observation of multiple pairs of players, over each of the time periods observed here. Since the ENO of some of the models increases over time (like equilibrium) while the other models predict somewhat less well over time, the simplest comparison with which to illustrate the method is to consider how well each model predicts over all 500 periods. For this purpose, observation of one pair of players provides a better prediction than equilibrium, but all of the other models are more informative than observing five pairs of players, and the reinforcement learning models and normalized fictitious play provide better predictions than observing more than ten pairs of players.

Evaluation of the effect of experience on the estimated ENO of the different models reminds us that the models are only approximations. For example, impulse balance equilibrium is much more useful in capturing the last 100 trials (ENO of 31.1) than the first 100 trials (ENO of 4.74). The low initial ENO suggests that (in violation of the implicit simplification assumption) behavior does not approach the impulse balance equilibrium immediately. Similarly, the large increase in the ENO of NFP (ENO of 7.1 and 107.1 in the first and last 100 trials, respectively) can be explained as a product of the simplification assumption. Specifically, to reduce the number of free parameters, this model assumes that the same weighting parameter (w) determines the adjustment of the propensities and the normalization factor $S(t)$. Relaxation of this assumption implies slower adjustment of the normalization factor and high ENO in the first as well as the last 100 trials.

The ENO values estimated above capture the generality of the current models in the context of the present experimental environment of repeated play of randomly selected binary constant sum games. To get some idea of the generality of

the results beyond this class of games we rely on the fact that the parameters of two models examined here, NFP and impulse balance equilibrium, were estimated in previous research that examined behavior in different settings. NFP, for example, was estimated in Ert and Erev (2007) to fit individual choice behavior in a setting with multiple alternatives with minimal prior information. Using the parameters estimated in this previous research, Table 2 shows the MSE scores and ENO of these models on the current data. The results show that in both cases the predictions of the models with these previously estimated parameters are less accurate (have higher MSE scores) than the predictions with the parameters estimated to fit the 30 game sample.

Nonetheless, both models have relatively high ENO using the parameter estimates from these different environments. For instance, over 500 trials the ENO of NFP with the parameters estimated to fit Ert and Erev's individual decisions data is 9.92.

4 Concluding remarks

There is widespread acceptance among economists [going back at least to Friedman's (1953) "as if" argument] that many economic theories of human behavior are best understood as approximations rather than as literal descriptions of human decision making. However, the experimental designs and statistical tests conducted to distinguish among models remain overwhelmingly oriented towards hypothesis testing. This may not be unreasonable when the only available data are noisy and sparse, but the increased use of laboratory experiments to test economic theories means that it will now be possible to gather sufficient data to reject theories of behavior that are not exactly correct. But the observation that sufficiently large data sets may allow any theory to be rejected cannot be allowed to immunize theories from empirical examination.

The present paper takes a step towards formalizing the discussion of the usefulness of approximation. By looking at a random sample of games, rather than a single game or sample of games selected by the experimenter, we are able to open a window on how well different theories predict on average. By defining the ENO of a theory as the number of observations that are needed to derive a better prediction, we are able to revisit the venerable debate among and between psychologists and economists on the merits of equilibrium and learning models of behavior in two person zero sum games. All of these theories can be, and have been, rejected; all of them can be distinguished from the process that actually generates the data. What the ENO tells us is how much data we would need before we could generate a better prediction than each theory.

Viewed in this way, we see that, after the players have gained sufficient experience, the mixed strategy equilibrium prediction for a game is "close enough" to be superior to the opportunity to observe a pair of players playing the game. But in the range of our experiment (500 periods) equilibrium did not (yet) yield a prediction superior to the observation of two pairs of players. Depending on the cost of experimentation on a previously unexamined game, and on the cost of error, this kind of result can guide the decision of whether to experiment.

Table 2 shows that variants of stochastic fictitious play predicted substantially better than equilibrium over the 500 periods studied, which is encouraging given

that the games studied here, two person zero sum games played under full information, are precisely the environment that motivated the study of fictitious play as a model of equilibration. (The fact that the most successful of the models of fictitious play we examine has an ENO of over 100 for the last 100 periods of play suggests that it may capture long term equilibration reasonably well.) Reinforcement learning did as well over all periods of play, which is not too surprising given the robust performance of related reinforcement learning models in a wide variety of settings.¹⁴ This success occurs even though reinforcement learning does not make use of all the information (about the other player's behavior) that is available in the experiment reported here. And, for the first 100 periods of play, reinforcement learning models outperformed the other models we considered, with an ENO of 37, suggesting that the early periods of learning also can be predicted with some success. And while more highly parameterized models can be used to fit the data more closely ex post, the ones studied here did not predict better than the simple models.

As economists are increasingly called on to evaluate and even to design new economic environments (cf. e.g., Roth 2002), the need for predictively powerful theories will grow in microeconomics generally, and in game theory in particular. The effective design of novel environments requires the ability to predict behavior without benefit of previous data from the same strategic environment. Useful approximations, if they can be identified, will be of enormous importance, even in the absence of theories that are exactly true. This paper is meant as a step towards evaluating the predictive value of approximations. The results reported here, particularly the results concerning reinforcement learning, leave us optimistic that quite general models, that may be applied in a wide variety of environments (in which agents may sometimes have only little information about the environment) have the potential to be useful approximations.

Appendix 1

a. EWA

EWA uses the same choice rule as SFP, and the following propensity adjustment rule:

$$q_k(t+1) = \{\phi \cdot N(t-1) \cdot q_k(t) + [\delta + (1-\delta) \cdot I(t,k)] \cdot v_k(t)\} / [N(t)].$$

where ϕ is a forgetting parameter, $N(1)$ is a free parameter, $N(t) = \rho N(t-1) + 1$ (for $t > 1$) is a function of the number of trials, ρ is a depreciation rate parameter, δ is a parameter that determines the relative weight for obtained and forgone payoffs, $I(t,k)$ is an index function that returns the value 1 if strategy k was selected in trial t and 0 otherwise, and $v_k(t)$ is the payoff that the player would receive for a choice of strategy k at trial t .

b. Impulse balance equilibrium (IBE)

IBE is defined on a transformed game. The computation of the transformed game includes three steps:

¹⁴ These settings include but are not limited to examples of bargaining and market games (Roth and Erev 1995), coordination (Erev and Rapoport 1998) team and public goods games (Bornstein et al. 1996; Nagel and Tang 1998), the repeated games of incomplete information (Feltovich 2000), signaling games (Cooper et al. 1997) and signal detection games (Erev 1998). See also Bereby-Meyer and Roth (2006).

- (1) Organization of the strategies to insure a clockwise best reply cycle. For example, Game 1 is replaced with Game 1'.

Game 1	Player 2		Game 1'	Player 2		Game 1''	Player 2	
Player 1	A	B	Player 1	A	B	Player 1	A	B
A	58,42	94,6	A	98,2	51,49	A	78,2	51,27.5
B	98,2	51,49	B	58,42	94,6	B	58,24	76,6

- (2) Computation of the maximal payoff that each player can obtain with certainty. In Game 1 this value, referred to as the aspiration level (*AL*), is 58 for Player 1, and 6 for Player 2.
- (3) Transformation of the payoffs to reflect under-sensitivity to gain relative to the reference point. Value *X* above the reference point is replaced with $X' = (X - AL)/\phi$ where ϕ captures the sensitivity to gain. The original model sets $\phi = 2$. The current analysis treats ϕ as a free parameter. Game 1'' is the transformed variant of Game 1 with $\phi = 2$.

Let $u(i, j, k)$ denote the payoff (in Game 1'') of player *i* when Player 1 selects *j* and player 2 selects *k*. Choice probabilities are determined as follows:

$$\begin{aligned}
 c &= [u(1, A,A) - u(1, B,A)]/[u(1, B,B) - u(1, A,B)] \\
 d &= [u(2, A,B) - u(2, A,A)]/[u(2, B,A) - u(2, B,B)] \\
 U &= (c/d)^{0.5} \\
 V &= 1/[(c * d)^{0.5}] \\
 P(\text{Player 1 selects A}) &= U/(1 + U) \\
 P(\text{Player 2 selects A}) &= V/(1 + V)
 \end{aligned}$$

Appendix 2

a. The derivation of ENO

Equal weight is implied when $MSE(\bar{X}_o) = MSE(G)$.

When there are *n* observations in each of *N* conditions,

$$MSE(\hat{\bar{X}}_o) = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{j=1}^n (x_{ij} - \bar{X}_{oij})^2,$$

where x_{ij} is the choice of subject *j* in task *i*, and \bar{X}_{oij} is the mean of the “other *n* - 1 subjects” in this task. Notice that

$$\bar{X}_{oij} = \frac{n\bar{X}_i - x_{ij}}{n - 1},$$

where \bar{X}_i is the mean of all *n* subjects in task *i*. These equations imply

$$MSE(\bar{X}_o) = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{j=1}^n \left(x_{ij} - \frac{n\bar{X}_i - x_{ij}}{n - 1} \right)^2$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{j=1}^n \frac{n^2}{(n-1)} \frac{(\bar{X}_i - x_{ij})^2}{(n-1)} \\
&= \frac{n}{(n-1)} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^n \frac{(\bar{X}_i - x_{ij})^2}{(n-1)} = \frac{n}{(n-1)} S^2,
\end{aligned}$$

where S^2 is the pooled error variance.

$M = \text{MSE}(G)$ is the estimate of the mean squared distance between the general model's prediction and each observation:

$$M = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{j=1}^n (x_{ij} - G_i)^2.$$

Thus, equal weight for predicting the next observation based on $n - 1$ observations is therefore expected when

$$\frac{n}{(n-1)} S^2 = M$$

and, the implied value of $n - 1$ (the size of the experiment expected to yield prediction as accurate as the model) is $S^2/(M - S^2) = \text{ENO}$.

b. The optimal weighting as a function of ENO

When $CD(\bar{X}_o, G) = S^2$:

$$\hat{W} = \hat{\beta}_1 = \frac{\text{MSE}(\bar{X}_o) - S^2}{\text{MSE}(\bar{X}_o) + M - 2S^2}.$$

Replacing $\text{MSE}(\bar{X}_o)$ with $\frac{n}{(n-1)} S^2$ we get

$$\begin{aligned}
\hat{W} = \hat{\beta}_1 &= \frac{\frac{n}{(n-1)} S^2 - S^2}{\frac{n}{(n-1)} S^2 + M - 2S^2} = \frac{S^2/n - 1}{S^2/n - 1 + M - S^2} \\
&= \frac{S^2}{S^2 + (n-1)(M - S^2)}, \\
\frac{1}{\hat{W}} &= \frac{S^2 + (n-1)(M - S^2)}{S^2} = 1 + (n-1) \frac{M - S^2}{S^2} = 1 + \frac{(n-1)}{\text{ENO}} \\
&= \frac{(n-1) + \text{ENO}}{\text{ENO}}.
\end{aligned}$$

Thus,

$$\hat{W} = \hat{\beta}_1 = \frac{\text{ENO}}{(n-1) + \text{ENO}}.$$

c. The relationship of ENO to the t statistic

The definition of the t statistic implies

$$t^2 = \frac{(\mu - \bar{X})^2}{S^2/n}.$$

Under the null hypothesis $\mu = 0$ (the model $G = 0$):

$$t^2 = \frac{\bar{X}^2}{S^2/n}.$$

The definition of S^2 implies

$$S^2 = \frac{\sum_{j=1}^n (\bar{X} - x_j)^2}{(n-1)} = \frac{\sum_{j=1}^n (\bar{X}^2 - 2\bar{X}x_j + x_j^2)}{(n-1)} = \frac{\sum_{j=1}^n (x_j^2) - n\bar{X}^2}{(n-1)}.$$

When the model states $G = 0$, and we consider only one task ($N = 1$) $nM = \sum_{j=1}^n (x_j^2)$. Thus,

$$S^2 = \frac{nM - n\bar{X}^2}{(n-1)},$$

$$\bar{X}^2 = M - \frac{n-1}{n}S^2 = M - S^2 + S^2/n.$$

Replacing for \bar{X}^2 in the definition of t^2 we get

$$t^2 = \frac{M - S^2 + S^2/n}{S^2/n} = \frac{M - S^2}{S^2/n} + 1 = \frac{n}{\text{ENO}} + 1.$$

References

- Allais, M.: Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'école américaine. *Econometrica* **21**, 503–546 (1953)
- Bereby-Meyer, Y., Alvin E. Roth: Learning in noisy games: partial reinforcement and the sustainability of cooperation. *Am. Econ. Rev.* **96**(4), 1029–1042 (2006)
- Bolton, G.E., Ockenfels, A.: ERC: a theory of equity, reciprocity and competition. *Am Econ Rev* **90**, 166–193 (2000)
- Bornstein, G., Winter, E., Goren, H.: Experimental study of repeated team games. *Eur J Polit Econ* **12**, 629–639 (1996)
- Brown, J.N., Rosenthal R.W.: Testing the minimax hypothesis: a re-examination of O'Neill's game experiment. *Econometrica* **58**, 1065–1081 (1990)
- Busemeyer, J.R., Wang, Y.M.: Model comparisons and model selections based on generalization criterion methodology. *J Math Psychol* **44**(1), 171–189 (2000)
- Bush, R.R., Mosteller, F.: *Stochastic Models for Learning*. New York: Wiley (1955)
- Camerer, C.F., Ho, T.: Experience-weighted attraction in games. *Econometrica* **67**, 827–874 (1999)
- Cheung, Y.W., Friedman, D.: Individual learning in normal form games: some laboratory results. *Games Econ Behavior* **19**, 46–76 (1997)
- Cohen, J.: The earth is round ($p < 0.05$). *Am Psychol* **49**(12), 997–1003 (1994)

- Cooper, D., Garvin, S., Kagel, J.: Signalling and adaptive learning in an entry limit pricing game. *Rand J Econ* **28**(4), 662–683 (1997)
- Edwards, W.: Probability learning in 1000 trials. *J Exp Psychol* **62**, 385–394 (1961)
- Engle-Warnick, J., Slonim, R.L.: Inferring repeated-game strategies from actions: evidence from trust game experiments. *Econ Theory* **28**(3), 603–632 (2006)
- Erev, I.: Signal detection by human observers: a cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychol Rev* **105**(2), 280–311 (1998)
- Erev, I.: SAS programs for the parameter estimation and ENO computation of the NFP model. Retrieved September 4, 2006, from the following web address: <http://iew3.technion.ac.il/Home/Users/erev.html#part2>
- Erev, I., Barron, G.: On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychol Rev* **112**, 912–931 (2005)
- Erev, I., Haruvy, E.: On the potential uses and current limitations of data driven learning models. *J Math Psychol* **49**(5), 357–371 (2005)
- Erev, I., Rapoport, A.: Magic, reinforcement learning and coordination in a market entry game. *Games Econ Behav* **23**, 280–298 (1998)
- Erev, I., Roth, A.E.: Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am Econ Rev* **88**(4), 848–881 (1998)
- Erev, I., Bereby-Meyer, Y., Roth, A.E.: The effect of adding a constant to all payoffs: experimental investigation, and implications for reinforcement learning models. *J Econ Organ Behav* **39**, 111–128 (1999)
- Erev, I., Roth, A.E., Slonim, S.L., Barron, G.: Predictive value and the usefulness of game theoretic models. *Int J Forecast* **18**(3), 359–368 (2002)
- Erev, I., Roth, A.E., Slonim, S.L., Barron, G.: Descriptive models as generators of forecasts with known equivalent number of observations (ENO). Working paper (2005)
- Ert, E., Erev, I.: Replicated alternatives and the role of confusion, chasing, and regret in decisions from experience. *J Behav Decis Making* (in press) (2007)
- Estes, W.K.: Toward a statistical theory of learning. *Psychol Rev* **57**, 94–107 (1950)
- Feltovich, N.: Learning and equilibrium in an asymmetric information game: an experimental study. *Econometrica* **68**(3), 605–642 (2000)
- Friedman, M.: *Essays in Positive Economics*. pp. 3–43 Chicago: University of Chicago Press (1953)
- Fudenberg, D., Levine, D.: *Theory of Learning in Games*. Cambridge: MIT Press (1998)
- Goeree, J.K., Holt, C.A.: Stochastic game theory: for playing games not just for doing theory. *Proc Nat Acad Sci* **96**, 10564–10567 (1999)
- Granger, C.W.J., Ramanathan, R.: Improved methods of combining forecasts. *J Forecast* **3**, 197–204 (1984)
- Gupta, S., Wilton P.C.: Combination of forecasts: an extension. *Manage Sci* **33**(3), 356–372 (1987)
- Hopkins, E.: Two competing models of learning in games. *Econometrica* **70**, 2141–2166 (2002)
- Luce, R.D.: *Individual Choice Behavior*. New York: Wiley (1959)
- Kagel, J.H., Roth, A.E.: *Handbook of Experimental Economics*. Princeton: Princeton University Press (1995)
- Nagel, R., Tang, F-F.: An experimental study on the centipede game in normal form—an investigation on learning. *J. Math. Psychol.* **42**(2), 256–284 (1998)
- Nash, J.F.: Equilibrium points in n -person games. *Proc Nat Acad Sci* **36**, 48–49 (1950)
- Ochs, J., Roth, A.E.: An experimental study of sequential bargaining. *Am Econ Rev* **79**, 355–384 (1989)
- O’Neill, B.: Nonmetric test of the minimax theory of two-person zero-sum games. *Proc Nat Acad Sci USA* **8**, 2106–2109 (1987)
- O’Neill, B.: Comments on Brown and Rosenthal’s reexamination. *Econometrica* **59**, 503–507 (1991)
- Roth, A.E.: Individual rationality as a useful approximation: comments on Tversky’s rational theory and constructive choice. In: Arrow, K., Colomatto, E., Perlman, M., Schmidt, C., (eds.) *The Rational Foundations of Economic Behavior*, pp. 198–202 New York: Macmillan, <http://kuznets.fas.harvard.edu/~aroth/rational.html> (1996)
- Roth, A.E.: The economist as engineer: game theory, experimental economics and computation as tools of design economics. *Econometrica* **70**(4), 1341–1378 (2002)

-
- Roth, A.E., Erev, I.: Learning in extensive-form games: experimental data and simple dynamic models in intermediate term. *Games Econ Behav, Spec Issue Nobel Symp* **8**, 164–212 (1995)
- Roth, A.E., Malouf, M.K.: Game theoretical models and the role of information in bargaining. *Psychol Rev* **86**, 574–594 (1979)
- Salmon, T.C.: An evaluation of econometric models of adaptive learning. *Econometrica* **69**, 1597–1628 (2001)
- Sarin, R., Vahid, F.: Predicting how people play games: a simple dynamic model of choice. *Games Econ Behav* **34**, 104–122 (2001)
- Selten, R., Chmura, T.: Stationary concepts for experimental 2×2 games. *Bonn Econ Discussion Paper* (2005)
- Siegel, S., Siegel, S., Andrews, J.M.: *Choice, Strategy, and Utility*. New York: McGraw-Hill (1964)
- Slonim, R., Roth, A., Erev, I.: Regularities in play of zero sum games with unique mixed strategy equilibria. *Working Paper* (2006)
- Suppes, P., Atkinson, R.C.: *Markov Learning Models for Multiperson Interactions*. Stanford: Stanford University Press (1960)
- Tversky, A.: Rational theory and constructive choice. In: Arrow, K., Colomatto, E., Perlman, M., Schmidt, C. (eds.) *The Rational Foundations of Economic Behavior*. pp. 185–197 New York: Macmillan, (1996)
- von Neumann, J.: Zur theorie der Gesellschaftespiele. *Math Ann* **100**, 295–320 (1928)
- Weibull, J.W.: Testing game theory, Stockholm School of Economics. *Working Paper* 382 (2000)
- Weibull, J.W.: Testing game theory. In: Huck, S. (ed.) *Advances in Understanding Strategic Behaviour: Game Theory, Experiments and Bounded Rationality. Essay in Honour of Werner Güth*. pp. 85–104 Palgrave: MacMillan, (2004)
- Wooders, J., Shachat, J.M.: On the irrelevance of risk attitudes in repeated two-outcome Games. *Games Econ Behav* **34**(2), 342–363 (2001)