International Journal of Forecasting 18 (2002) 359–368

international journal of forecasting

www.elsevier.com/locate/ijforecast

Predictive value and the usefulness of game theoretic models

Ido Erev[a,d,*], Alvin E. Roth[b], Robert L. Slonim[c], Greg Barron[d]

[a]*Graduate School of Business, Columbia University, New York, NY, USA*
[b]*Harvard University, Cambridge, CN, USA*
[c]*Case Western Reserve University, Cleveland, OH, USA*
[d]*Technion, Haifa, Israel*

Abstract

Green [*Int. J. Forecasting* (2002)] reports that in certain settings predictions made by game theorists can be outperformed by the outcome of a short role playing exercise. Goodwin [*Int. J. Forecasting* (2002)] argues that this does not imply that game theoretic analysis cannot be useful. The current paper discusses two types of observations that support this assertion. First, there are many important settings in which game theoretic models have high forecasting power. Two examples: the aggregate outcome of entry job markets, and the outcome of repeated interactions are summarized here. The second observation concerns the possibility of objectively forecasting the predictive value of specific models (and methods) on particular domains. To increase our understanding of the value of role playing, we suggest that future research focus on estimating the predictive value of this method using a random selection of problems from a well defined set. © 2002 International Institute of Forecasters. Published by Elsevier Science B.V. All rights reserved.

Keywords: Predictive value; Role playing

1. Introduction

Green (2002) demonstrates that predictions obtained from a short role playing exercise (see Armstrong, 2001) can be better than those solicited by email from experienced game theorists. The current paper builds on Goodwin's (2002) reply, and argues that these results do not mean that game theory cannot be useful. Indeed, there are relatively well defined sets of situations in which previous empirical and experimental research has documented the high predictive value of specific game theoretic models. Two examples (bargaining games and auctions) are discussed in Bolton (2002) and Shefrin (2002). The current paper presents two additional examples and a method that can be used to forecast the predictive value of a specific model.

The first example we discuss involves the successful design and implementation of large markets. It is virtually impossible to predict specific market outcomes, since the tastes and opportunities of market participants are largely unobservable. But it is nevertheless possible to

*Corresponding author.
E-mail address: erev@tx.technion.ac.il (I. Erev).

0169-2070/02/$ – see front matter © 2002 International Institute of Forecasters. Published by Elsevier Science B.V. All rights reserved.
PII: S0169-2070(02)00023-7

predict what kinds of rules of market organization will lead to orderly markets. This kind of work has led to successful designs of large markets, such as the entry level labor market for American doctors (the National Resident Matching Program), and the auction markets for radio spectrum conducted by the U.S. Federal Communications Commission (and similar markets in other countries) considered by Shefrin (2002).

The second example is concerned with the development of game theoretic models with higher forecasting ability for smaller-scale strategic environments, in the short and intermediate term. Most of classical game theoretic analysis concerns modeling the long term 'equilibrium' behavior in a game, but recent attention to models of learning in games suggests that these have better forecasting ability in the near term.

The fourth section of this paper formulates an approach for measuring the forecasting value of game theoretic (and other) models (and methods) in specific domains. It shows how experimental procedures that involve random sampling of problems can be used to forecast predictive value. We believe that methods of this type are needed to understand the relative value of game theoretic models and role playing methods in different settings.

2. Market design

Game theory is concerned with the 'rules of the game,' and as such it is the part of economic theory that has come to the fore in understanding how a market's rules influence its operation. Game theoretic models have recently played a large role in market design. An example will illustrate how such models are useful.

One of the main functions of many markets and social processes is to match one kind of agent with another: e.g. students and colleges, workers and firms, marriageable men and women. A class of game-theoretic models of 'two-sided matching markets' for studying such processes was introduced by Gale and Shapley (1962). A market is two-sided if there are two sets of agents, and if an agent from one side of the market can be matched only with an agent from the other side. Gale and Shapley proposed that a matching could be regarded as stable only if it left no pair of agents on opposite sides of the market who were not matched to each other but would both prefer to be. They showed that a special property of two-sided (as opposed to one or three-sided) markets is that stable matchings always exist (at least when agents' preferences are uncomplicated).

However many markets, particularly entry level professional labor markets, experience market failures which prevent them from achieving stable matchings. In one common form of this market failure, the date of first appointment unravels in time, becoming earlier and earlier from year to year (see e.g. Roth and Xing (1994) for a description of several dozen markets and submarkets that have experienced unraveling at some point in their history). In the United States, two markets that are presently experiencing this kind of unraveling are the market for law clerks for Federal appellate judges (cf. Avery, Jolls, Posner, & Roth, 2001), in which offers are presently made almost two years in advance of employment, and the market for college admissions (cf. Avery, Fairbanks, & Zeckhauser, 2001), in which elite colleges admit a high percentage of their entering classes through 'early decision' programs. Both of these are 'gateway' markets, whose outcomes have a strong influence on future career paths.

Quite a few markets have resolved such market failures by attempting to organize centralized clearinghouses, which would permit most positions in the market to be filled at the same time. For example, such clearinghouses have been organized in medical and other health-care markets in the U.S., Britain, and Canada, and in regional markets for new lawyers in Canada. Some of these clearinghouses have succeeded, while others have failed. The best predictor of success turns out to be whether the market produces stable matchings (cf. Roth, 1984, 1990, 1991; Roth & Xing, 1994).

As market conditions change, some of these clearinghouses need to be redesigned in order to continue to produce stable matchings. For example, the entry level market for American physicians, which fills roughly 20 000 positions per year, now has to deal with perhaps 1000 applicants per year who go through the match as married couples and desire two positions in the same city, and more than twice that many applicants who need to be placed into two consecutive positions in order to qualify for a particular medical specialty. Game theory was at the heart of the recent redesign of the medical clearinghouse (the National Resident Matching Program), as described in Roth and Peranson (1999). The Roth–Peranson design has also been adopted by entry level labor markets in other professions since its adoption by American physicians.[1] Note that the game-theoretic idea of stability plays a critical role in the success of such markets, but that what constitutes a stable outcome depends entirely on the preferences of the market participants. So the task of the market designer is to formulate rules that give firms and workers the incentive to participate in the market in an orderly way.

It should be emphasized that, in this kind of application as in others, predictions made using game theory are not made entirely a priori. The formal theory needs to be validated by empirical observation, both in the field and in the laboratory (see e.g. Kagel and Roth (2000) for an experiment in support of this kind of design effort). This is probably going to be fairly typical of the use of game theory in design: for a review that also includes discussion of the recent design of national auctions for radio spectrum licenses, see Roth (2002).

3. Repeated interactions

Recent research demonstrates that in many settings the development of game theoretic models with high forecasting value requires a relaxation of the high rationality assumption (made in classical game theoretic models). One example involves the modeling of the effect of repeated interactions. To demonstrate this line of research the current section focuses on the repeated play of matrix games with unique mixed strategy equilibrium in which players cannot reciprocate. In a typical study two players participate in a multi trial play of the game. In each trial Player 1 chooses between the rows of a matrix and Player 2 chooses between the columns. The choices determine which cell of the payoff matrix will be used to compute the trial's payoffs. Review of all the published experiments studying games of this type (Erev & Roth, 1998) reveals that the classical game theoretic prediction (maximin, mixed strategy equilibrium) has very low predictive value. Indeed, the mean squared deviation (MSD) between the observed and predicted choice-proportions was larger than the MSD score expected under the prediction of random behavior.

However, Erev and Roth (and see Sarin & Vahid, 2001; Erev, Bereby-Meyer, & Roth, 1999) found that relatively small modifications of the game theoretic model can dramatically

[1] Other markets that have adopted it to date are, in the United States, Postdoctoral Dental Residencies, Osteopathic Internships, Osteopathic Orthopedic Surgery Residencies, Pharmacy Practice Residencies, and Clinical Psychology Internships, and, in Canada, Articling Positions with Law Firms in Ontario, Articling Positions with Law Firms in Alberta, and Medical Residencies.

improve forecasting power. The classical game theoretic prediction is based on the assumption that each player plays best reply to the strategy of the other players. In the models considered by Erev and Roth each player slowly learns to adjust their behavior to the previously obtained outcomes. This modification does not necessarily change the long-term predictions of the model (because the learning process can lead to equilibrium), but it appears to provide much better forecasting of the first hundred trials. Interestingly, the good forecasting value was obtained even when the models' parameters (one to three) were fixed over the 12 games, i.e. this is not a case in which the parameters have to be fitted to the games, but rather a case in which a general model has predictive power over the whole set of games studied.

4. Forecasting predictive value

To evaluate the implications of a particular model (like the adaptive game theoretic models discussed above) and/or a method (like role playing) in a particular setting it is necessary to address the 'second order forecasting problem.' That is, the forecasting of the predictive value of the model and/or method. This problem can be addressed by considering a well defined universe of situations and asking how well the model/method will predict, on average, behavior in a situation randomly sampled from that universe. In the current section we consider how experimental methods can be used to formalize and assess this kind of predictive value.

For illustrative purposes, we concentrate here on the simple universe of two-person zero sum games with a unique mixed strategy equilibrium, in which each player has only two strategies, and payoffs are binary lotteries, i.e. payoffs are the probability that each player will get a fixed prize (this universe is a subset of the situations considered in Section 3). The leftmost column of Fig. 1 presents ten randomly generated games of this type. In Fig. 1's examples only two payoffs are possible (0 or 1). The selected cell determines the probability of the higher payoff for each player. For example, when cell (A1, B2) is selected in the game presented in the bottom row of Fig. 1, the probability that Player 1 wins 1 is 0.74 and the probability that Player 2 wins 1 is 0.26. The classical game theoretic (maximin, mixed strategy equilibrium) prediction for this game is that Player 1 will play A1 with probability 0.985 and Player 2 will play A2 with probability 0.794.

For example, consider the predictions of the proportion of A1 choices in trials 401–500 of the games summarized in Fig. 1. The true proportion can be predicted with a specific model and with the results of an experiment with k pairs of subjects. Obviously the accuracy of the experiment increases with k. The model's Predictive Value (PV) is defined as the size of the experiment (the value of k) that has to be run to provide predictions that are as accurate as the model's prediction.

Accuracy is defined here as the mean squared deviation (MSD) between the prediction and the observation. Thus, to estimate the PV of a model, we have to estimate the MSD between the model's prediction, Pr, and the observations, and compare that to the MSD between the prediction derived from subsamples of size k and observations from outside the subsample. Formally, denote by $M_{\mathrm{Pr},n}$ the mean distance between the prediction Pr and the n observations,

$$M_{\mathrm{Pr},n} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mathrm{Pr})^2,$$

and denote by $M_{k,n}$ the mean squared distance between the mean of each subsample of size k and each observation x not in the subsample, for all the s sub-samples of size k,

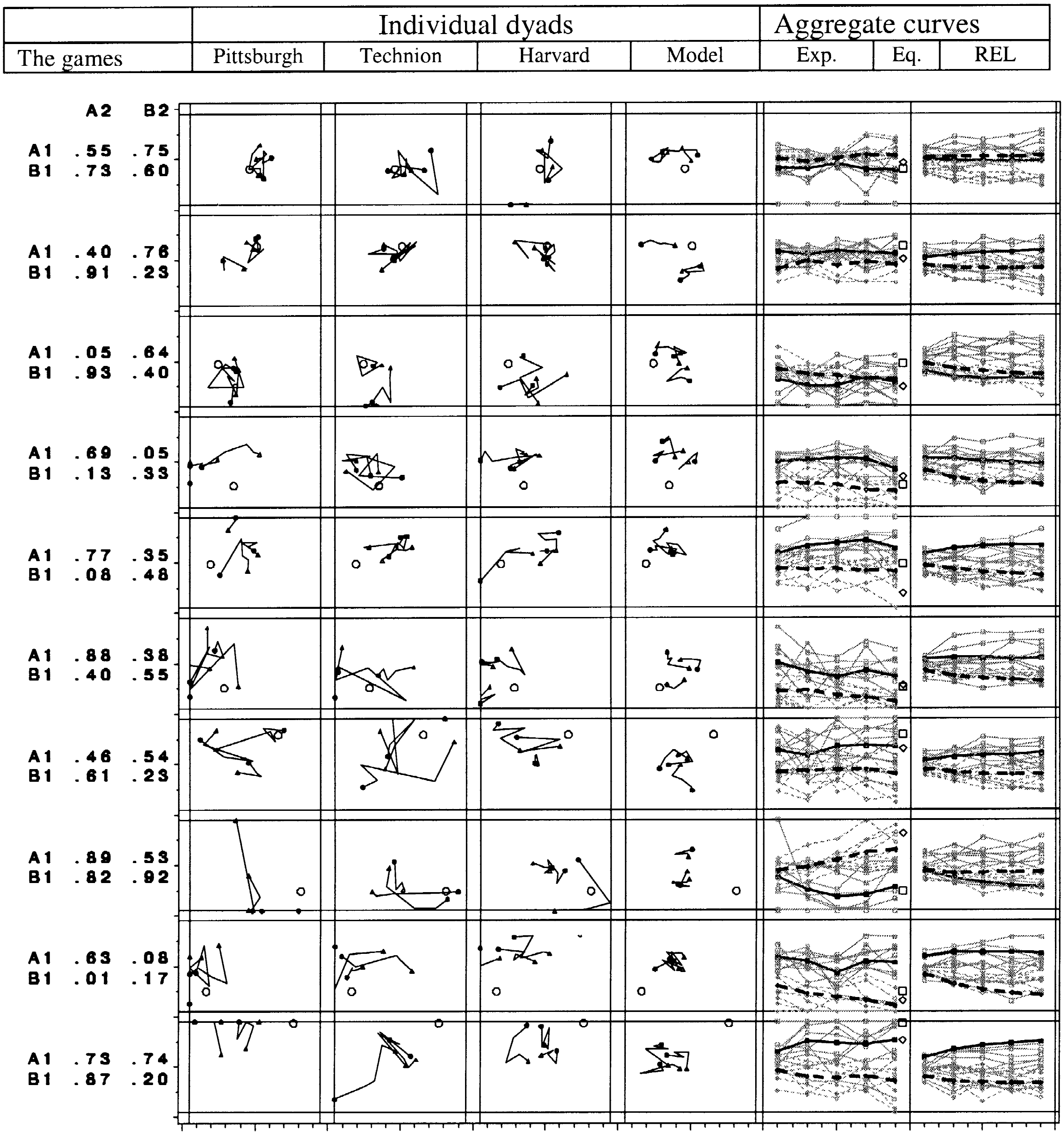

Fig. 1. Individual dyads: Each cell on the left shows three dyads observed at a different university. Each curve shows the choice probabilities in 5 blocks of 100 trials. The Y-axis (0 to 1) is P(A1) and the X-axis (0 to 1) is P(A2). The Δ represents the first block and the · represent the last. The O is the equilibrium. The right hand columns show the choice probabilities as a function of time (5 block of 100 trials). The light curves shows nine dyads and the dark curves show the means. Player 1's curves are solid with squares, Player 2's curves are dashed with diamonds. The equilibrium predictions are found at the right of the second to the last column.

$$M_{k,n} = \frac{1}{s} \sum_{j=1}^{s} \left(\frac{1}{n-k} \sum_{i=1, i \notin j}^{n} (x_i - \bar{x}_j)^2 \right)$$

where s is the number of subsets of observations of size k in the sample of n observations ($s = n!/[k!(n-k)!]$ and j is an index for each subset.

Simple algebra (see Barron, 2000) reveals that

$$M_{1,n} = 2\sum(x_i - \bar{x}_j)^2/(n-1) = 2S^2$$

where S^2 is the familiar sample variance and,

$$M_{k,n} = M_{1,n}\left(\frac{k+1}{2k}\right)$$

As k increases, $M_{k,n}$ decreases and approaches $M_{1,n}/2$ from above as k (and n) go to infinity. Thus, when $M_{\mathrm{Pr},n} < M_{1,n}/2$ the data suggest that no experiment will outperform the model (the model's PV is very large). When $M_{\mathrm{Pr},n} > M_{1,n}/2$ the model's PV can be precisely calculated as the k^* that solves the equation

$$M_{\mathrm{Pr},n} = M_{1,n}\left(\frac{k^*+1}{2k^*}\right)$$

The solution is:

$$PV_n = k^* = \frac{M_{1,n}/2}{M_{\mathrm{Pr},n} - M_{1,n}/2}$$

In Erev, Roth, Slonim, and Barron (2002) we show that PV is closely related to common statistics (including constrained regressions, Minimum variance weight in Bayesian statistics, Student t, Cohen's d). Thus, like these other statistics the PV of a particular model in a particular domain can be estimated by taking a random sample of observations (prediction tasks) from the relevant population of tasks.

To allow robust estimate (and maintain the relationship to regression analysis) it is constructive to estimate the *pooled* PV. Pooled PV in a population of tasks is defined here as the number of subjects (observations) needed (under the constraint of the same number in each game) to obtain an MSD score (over games) that is identical to the MSD score of the model (over the games).

4.1. Experiment

To demonstrate the use of the PV measure, we present an experiment designed to compute the PV of two game theoretic models in predicting behavior in two-person constant-sum games of the type presented in Fig. 1. The experiment studies behavior in 10 randomly selected games of this type (cf. Fig. 1). Each subject played 500 repetitions of one of the games against a fixed, anonymous opponent.

As noted above, the numbers in each matrix represent probabilities that the players will win a fixed amount w on each trial. For example, if on a given trial both players choose action 'A,' then player 1 will win w with the specified probability $p1$, and player 2 will win w with probability $1 - p1$. A player who does not win w earns zero for that period. In each of the games played, w 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). 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 $p1$ through $p4$ were independently chosen from the uniform distribution on the values [0.00, 0.01, . . . 0.99, 1.00]. Games generated in this way were included in the sample if they had a unique mixed strategy equilibrium.

The participants were informed that they are playing constant sum games (see instructions in http://techunix.technion.ac.il/~barron) but did not know the probabilities that defined the game they were playing. After each period of play each player learned whether or not they received the payoff w ($w = \$0.04$), but did not know whether the other player received w.

Because the games have binary lottery payoffs, the equilibrium predictions can be determined without estimating any unobservable parameters (involving risk aversion) (Kagel & Roth, 1995; Roth & Malouf, 1979; Wooders & Shachat, 2001).[2] So a single random sample of games would be adequate for measuring the closeness of the equilibrium prediction to the observed behavior. But the learning models have free parameters, which must be estimated. Since we are interested in predictive power for new games, we collect data from two distinct random samples of games, so the parameters of the learning model can be estimated from one sample, and used to predict behavior in the other sample.

Each game was played by nine pairs, three each in Boston, Haifa, and Pittsburgh (at the experimental laboratories of Harvard, Technion, and University of Pittsburgh). That is, although the games are randomly generated, the players are not, but we compare players drawn from different subject pools to make sure that the behavior we observe is robust to choice of subject pool.

4.2. Two models:

We will consider the predictive value of two models: equilibrium, and the reinforcement learning (REL) model proposed by Erev et al. (1999). 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).[3]

The REL model assumes that the probability of selecting alternative k at trial t is given by:

$$P_k(t) = \frac{e^{q_k(t)\lambda/S(t)}}{\sum_{j=1}^{2} e^{q_j(t)\lambda/S(t)}}$$

where $q_j(t)$ is the propensity to select strategy j and λ is a payoff sensitivity measure, and $S(t)$ is a measure of payoff variance. Updating occurs only for the selected strategy (j):

$$q_j(t+1) = [N(1) + C_j(t) - 1) \cdot q_j\,(t) + x]/[N(1) + C_j(t)]$$

where $C_k(t)$ is the number of times that strategy k was selected in the first t trials, x is the obtained payoff, and $N(1)$ is a parameter that determines the weight of the initial (uniform) tendencies.

$S(1)$ is the expected absolute distance between the payoff from random choices and the expected payoff given random choice ($A(1)$), and

[2] 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 (2000) about experimental control of other aspects of players' preferences are ameliorated.

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

$$S(t+1) = S(t)[t + 2N(1)]/[t + 2N(1) + 1] + |A(t) - x|/[t + 2N(1) + 1]$$

where

$$A(t+1) = A(t)[t + 2N(1)]/[t + 2N(1) + 1] + x[t + 2N(1) + 1]$$

Notice that the REL model has two free parameters λ and $N(1)$. The current analysis use the estimates that best fit the data examined by Erev et al.: $\lambda = 2.8$ and $N(1) = 30$.

4.3. The predictions and the observed behavior

The main experimental results are summarized in the first three and the fifth columns of Fig. 1 (columns four and six present the predictions of one of the REL model). Each game in the sample is displayed on the left, and the first three graphs show the behavior of each subject pair at the three universities in the three subject pools. The vertical axis shows for player 1, and the horizontal axis for player 2, the frequency with which each played his first strategy in a block of 100 observations. The first block of 100 (periods 1–100) are denoted by a triangle, and the last 100 (periods 401–500) are denoted by a dark oval. The equilibrium prediction for each game is denoted by a light oval, so the first three graphs allow the evolution of play by each of the nine pairs to be seen with respect to the equilibrium prediction.

The fifth column graphs the probabilities over time, both individual subject pairs (light lines), and the mean (dark lines) of all nine row players [squares], and column players [diamonds]. At the right of that column is the equilibrium prediction for each player for that game, so the graph makes clear how the means approach equilibrium over time. Looking at the individual pairs reminds us how much variance there is, and hence why prediction is difficult.

The fourth and sixth columns in Fig. 1 show the predictions of the REL model. The simulation predictions are displayed in the same format as the experimental data. The fourth column presents 3 randomly selected simulated pairs, and the sixth column shows the means (over 1000 simulations) and 9 individual simulations.

The top panel of Table 1 shows the MSD's of the two models. The lower panel shows the average sample variance $S^2 = M_{1,n}/2$ and PV for different statistics for each model. The results support the assertions made above. Whereas the classical equilibrium model has

Table 1

		Predicting goal			
		First 100	5 of 100	All 500	Last 100
MSD ($\times 100$)	Equilibrium	7.19	6.54	5.27	6.82
	REL	2.71	3.08	1.8	3.92
S^2		2.56	2.72	1.54	3.56
PV	Equilibrium	0.55	0.71	0.48	1.09
	REL	16.4	9.48	5.84	9.96

The top rows present the MSD scores of the two models ($\times 100$) for four distinct predicting goals. The lower rows show the variance (S^2) and the estimated PV.

very low PV (0.48–1.09), the assumption of slow adjustment (the REL model) has relatively high PV (5.84–16.4).

5. Discussion

Our confidence in the value of game theoretic analysis stems from two related observations. First, there are many important settings in which game theoretic models have high forecasting power. The current paper focuses on two examples: the aggregate level outcome of entry job markets, and the outcome of repeated interactions. Other important examples involve bargaining games (see Bolton, 2002) and auctions (see Shefrin, 2002). The second observation concerns the possibility of objectively forecasting the predictive value of specific models (and methods) on particular domains. This is important because the value of game theoretic (and other) models and methods is situation specific (see Goodwin, 2002). Using Bolger and Wright's (1994) terminology, the computation of predictive value can facilitate estimates of 'ecological validity.'

Because the value of game theoretic models is situation specific, it is quite possible that in the context of the complex conflicts examined by Green (2002) role playing based predictions will be more useful. Yet, to increase our understanding of the value of role playing, we suggest that future research will focus on estimating the predictive value of this method in a well defined set of situations. Ideally, the sampling of natural conflicts and the reduction of the sampled conflicts to role playing exercises should be performed before the observation of the outcomes of the conflict. The most difficult parts of prediction, not addressed in Green (2002), concern the prediction of outcomes *before* they happen, and therefore before they can inform the design of the role playing exercise.

In closing, the perfectly rational model of classical game theory, and even the simple adaptive learning models we have considered here, are not proposed as 'true descriptions' of the world, but rather as useful approximations. How useful is an approximation depends in part on the use to which it is being put, and can only be assessed empirically. We have proposed here a framework in which the predictive value of a theoretical model can be assessed in terms of how many empirical observations would be necessary to obtain a superior prediction.

Acknowledgements

This research was supported by a grant from NSF and the USA–Israel Binational Science Foundation.

References

Armstrong, J. S. (2001). Role playing: a method to forecast decisions. In Armstrong, J. S. (Ed.), *Principles of forecasting: a handbook for researchers and practitioners*. Norwell, MA: Kluwer Academic Publishers, pp. 25–30.

Avery, C., Fairbanks, A., & Zeckhauser, R. (2001). What worms for the early bird? Early admissions at selective colleges. Working paper, Kennedy School of Government, Harvard University.

Avery, C., Jolls, C., Posner, R. A., & Roth, A. E. (2001b). The market for federal judicial law clerks. *University of Chicago Law Review*, *68*(3), 793–902.

Barron, G. (2000). The effect of feedback on decision making under uncertainty and risk: the predictive value of descriptive models. Masters Thesis, Technion.

Bolger, F., & Wright, G. (1994). Assessing the quality of expert judgment: issues and analysis. *Decision Support Systems*, *11*, 1–24.

Bolton, G. E. (2002). Game theory's role in role playing. *International Journal of Forecasting*, *18*, 353–358.

Bolton, G. E., & Ockenfels, A. (2000). ERC: a theory of equity, reciprocity and competition. *American Economic Review*, *90*(1), 166–193.

Erev, I., Bereby-Meyer, Y., & Roth, A. E. (1999). The effect of adding a constant to all payoffs: experimental investigation, and implications for reinforcement learning models. *Journal of Economic and Organizational Behavior*, *39*, 111–118.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: reinforcement learning in experimental games with unique mixed strategy equilibria. *American Economic Review*, *88*(4), 848–881.

Erev, I., Roth, A. E., Slonim, R. L., & Barron, G. (2002). Combining theoretical predictions with experimental evidence: standardized approximations and their predictive value. Working paper.

Gale, D., & Shapley, L. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, *69*, 9–15.

Goodwin, P. (2002). Forecasting games: can game theory win. *International Journal of Forecasting*, *18*, 369–374.

Green, K. C. (2002). Forecasting decisions in conflict situations: a comparison of game theory, role playing and unaided judgment. *International Journal of Forecasting*, *18*, 321–344.

Kagel, J. H., & Roth, A. E. (1995). *Handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Kagel, J. H., & Roth, A. E. (2000). The dynamics of reorganization in matching markets: a laboratory experiment motivated by a natural experiment. *Quarterly Journal of Economics*, *115*(1), 201–235.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Science*, *36*, 48–49.

Ochs, J., & Roth, A. E. (1989). An experimental study of sequential bargaining. *American Economic Review*, *79*, 355–384.

Roth, A. E. (1984). The evolution of the labor market for medical interns and residents: a case study in game theory. *Journal of Political Economy*, *92*, 991–1016.

Roth, A. E. (1990). New physicians: a natural experiment in market organization. *Science*, *250*, 1524–1528.

Roth, A. E. (1991). A natural experiment in the organization of entry level labor markets: regional markets for new physicians and surgeons in the UK. *American Economic Review*, *81*, 415–440.

Roth, A. E. (2002). The economist as engineer: game theory, experimental economics and computation as tools of design economics. *Econometrica*, in press.

Roth, A. E., & Malouf, M. K. (1979). Game-theoretic models and the role of information in bargaining. *Psychological Review*, *86*, 574–594.

Roth, A. E., & Peranson, E. (1999). The redesign of the matching market for American physicians: some engineering aspects of economic design. *American Economic Review*, *89*(4), 748–780.

Roth, A. E., & Xing, X. (1994). Jumping the gun: imperfections and institutions related to the timing of market transactions. *American Economic Review*, *84*, 992–1044.

Sarin, R., & Vahid, F. (2001). Predicting how people play games: a simple dynamic model of choice. *Games and Economic Behavior*, *34*, 104–122.

Shefrin, H. (2002). Behavioral game theory. *International Journal of Forecasting*, *18*, 375–382.

Weibull, J. W. (2000). Testing Game Theory. Stockholm School of Economics, Working Paper 382.

Wooders, J., & Shachat, J. M. (2001). On the irrelevance of risk attitudes in repeated two-outcome games. *Games and Economic Behavior*, *34*(2), 342–363.

von Neumann, J. (1928). Zur theorie der Gesellschaftespiele. *Mathematische Annalen*, *100*, 295–320.