Int J Game Theory
DOI 10.1007/s00182-006-0015-8

ORIGINAL ARTICLE

Foregone with the wind: indirect payoff information and its implications for choice

Brit Grosskopf · Ido Erev · Eldad Yechiam

Received: 15 October 2004 / Accepted: 14 April 2006

Abstract Examination of the effect of information concerning foregone payoffs on choice behavior reveals a complex pattern. Depending on the environment, this information can facilitate or impair maximization. Our study of nine experimental tasks suggests that the complex pattern can be summarized with the assumption that initially people tend to be highly sensitive, and sometimes too sensitive, to recent foregone payoffs. However, over time, people can learn to adjust their sensitivity depending on the environment they are facing. The implications of this observation to models of human adaptation and to problems of mechanism design are discussed.

Keywords Foregone payoff · Adaptive behavior · Reinforcement learning · Fictitious Play · Directional learning · Bandit problems.

1 Introduction

In many situations people obtain payoff information only from the actions they have chosen. Diners often do not know how a different choice of main course

We thank Nick Feltovich, Ernan Haruvy, Rosemarie Nagel and Rajiv Sarin for helpful comments and suggestions. This research was initiated while the first author was visiting the Technion in Haifa. We very much appreciate the hospitality of the Minerva Center.

B. Grosskopf (✉)
Department of Economics, Texas A&M University,
4228 TAMU, College Station, TX 77843-4228, USA
e-mail: bgrosskopf@econmail.tamu.edu

I.Erev · E. Yechiam
Behavioral Science Area, Faculty of Industrial Engineering and Management,
Technion - Israel Institute of Technology,
Haifa 32000, Israel

or even restaurant would have turned out. Poker players usually do not know their foregone payoffs. However, in probably as many situations people obtain payoff information from actions they did choose *and* from those they did not. Investors in the stock market usually observe the performance of their mutual fund along with the performance of other funds. Roulette players in a casino directly observe what they would have gotten if they had placed their bets differently.[1]

In traditional, static choice models of decisions and games indirect payoff information does not play any role because agents choose their optimal actions given their beliefs. In models of learning indirect payoff information can be assumed to play a role. In the popular and widely used model of fictitious play (e.g., Brown 1951; Robinson 1951; Fudenberg and Levine 1998) people do not observe indirect payoffs, only the realized state of the world, which allows them to infer foregone payoff information given their knowledge of the payoff function. In the presence of foregone payoff information agents can still behave as in fictitious play if they give equal weight to the payoffs from chosen and unchosen actions, i.e., treat them symmetrically even without the knowledge of the payoff function (Cheung and Friedman 1997). Experience-weighted-attraction learning, however, treats them asymmetrically with foregone payoff information being given less weight (e.g., Camerer and Ho 1999).[2] Yet other classes of models do not take into account foregone payoff information (e.g., Roth and Erev 1995; Sarin and Vahid 1999), perhaps because such information is not assumed to be available.

In this paper, we conduct individual choice experiments with and without indirect payoff information. Our objective is to uncover any systematic effects of this foregone payoff information. We choose individual choice experiments so as not to confound the influence of multiple players simultaneously responding to such information. We view this as a natural starting point to study the effect of indirect payoff information in games.

To analyze the effect of feedback on human choice behavior the present research focuses on basic adaptation tasks, in which feedback concerning obtained and foregone payoffs is all the available information. In a typical experimental task the participants are recruited to "play the money machine game." They are presented with two (or more) unmarked alternatives (represented as buttons on a computer screen, c.f. Appendix) and told that in each trial of the study they should select one of the alternatives. Each choice results in a number of points that are added to the participant's total earning. The points are converted to money at the end of the experiment. Each of the experiments considered here contrasted two information conditions. In the "obtained only"

[1] Yet, there are other situations where foregone payoffs are partially available. Unsuccessful bidders in a first-price auction usually know how much they would have gotten if they had bid differently. Players going for the deer in a stag hunt game know what they would have gotten if they had gone for the rabbit.

[2] A more agnostic approach is taken by Heller and Sarin (2002) who allow for indirect payoff information to be inflated or deflated.

condition the available feedback was limited to the obtained payoff for the selected alternative. In the condition "with foregone" the participants received complete feedback about their obtained as well as about their foregone payoffs.

Notice that in the current settings almost any behavior can be justified as "rational"; it is possible to find a particular set of prior beliefs that would lead Bayesian agents to exhibit this behavior. Finding the optimal response in the current paradigms (and in the natural situations we hope to address) is more difficult than in bandit problems (see, Berry and Fristedt 1985) because the decision makers are not informed that the payoff distributions are stable. To address this observation the current research focuses on the behavioral regularities, and not on the differences between these regularities and the (unknown) prescriptions of the rationality assumption.

The current analysis reveals four distinct effects of foregone payoffs. Section 2 reviews four previously studied binary choice tasks, in which information regarding foregone payoffs affects sequential dependencies (e.g., changes from round-to-round behavior) but has almost no effect on aggregate behavior.

Section 3 presents a new experiment showing that when the payoffs of the different alternatives are positively correlated, information concerning foregone payoffs can have a large effect. In this case the additional information facilitates payoff maximization.

Section 4 shows that in a noisy multi-alternative environment that resembles the casino example, the addition of foregone payoff information can lead to an initial preference for risky options with low expected value. Interestingly, this negative effect diminishes with experience.

Section 5 demonstrates that the negative effect observed in section 4 is not due simply to the large number of alternatives. Indeed, when the environment is deterministic, the (positive) effect of foregone payoffs appears to increase with the number of alternatives.

In order to evaluate the predictability of the effect of foregone payoffs, section 6 tries to summarize the four sets of studies with general models with a single set of parameters. This analysis reveals that the different effects observed in the nine experimental tasks cannot be reproduced with the popular models proposed in previous research. Nevertheless, these models can be easily modified to capture observed behavior. For example, the main results can be reproduced with a fictitious play model that assumes diminishing exploration.

2 Aggregate and sequential effects of foregone payoffs

We start with revisiting four binary choice tasks analyzed in Haruvy and Erev (2001). They evaluated the effect of foregone payoff information in the following four problems.

Problem 1 H: (11 with certainty) L: (10 with certainty)
Problem 2 H: (1 with p=0.5; 21 otherwise) L: (10 with certainty)
Problem 3 H: (−10 with certainty) L: (−11 with certainty)
Problem 4 H: (−10 with certainty) L: (−1 with p=0.5; −21 otherwise)

The experiment consisted of 200 trials. In each trial the participants were asked to select between two unmarked buttons. One button led to a draw from distribution H (high expected value), and the other from distribution L (low expected value). Each subject only faced one problem under one feedback condition. Fifty-six subjects (14 per problem) were run in condition "Obtained Only" (OO). Forty subjects (10 per problem) were run in condition "With Foregone" (WF). The conversion rate was 0.01 Shekel (0.25 cent) per point. In problems 3 and 4 participants received a show-up fee of 40 Shekels, no show-up fee was given in problems 1 and 2.

The participants in this and all other studies reported here were undergraduates (mostly engineering students) at the Technion (Israel) recruited by campus advertisements. On average, participants in all of the presented problems took about 20 min to complete the 200 trials of the individual decision making task and earned roughly 24 Shekels (then about \$6).

The proportion of maximization (H choices) in 10 blocks of 20 trials each is presented in Fig. 1. The results reveal slightly higher maximization rates when foregone payoffs were provided (condition WF) in three of the four problems (problems 1, 2 and 3) although the differences are not significant.

Even though information about foregone payoffs does not affect maximization rates, it has a strong effect on individual choice behavior. Table 1 presents the observed probabilities of a switch in choices from trial t to trial $t+1$ in Problem 2 and 4 (with foregone) conditional on the observed outcome of the previous trial. The results show that in both problems subjects were more likely to change their choices when foregone payoffs were attractive, i.e., bigger than obtained payoffs.

The apparent inconsistency between the weak effect of the foregone on the aggregate and the clear effect on the switching probabilities can be explained with the assertion that the foregone payoff information is used, but its usage does not have a large effect on maximization. A natural abstraction of this idea is provided with learning models that assume a large recency effect. Models of this type imply a positive and a negative effect of foregone payoff information on maximization. A positive effect is predicted because foregone payoffs increase the information about the alternatives that is available to the decisions makers.

Table 1 Observed conditional probabilities of switches in choices from trial t to trial $t+1$ in problems 2 and 4 with foregone payoff information

				P(switch) given the event		
Problem	Choice at t	Possible obtained	Possible foregone	Foregone> observed	Foregone< observed	Difference
2	L	10	1 or 21	0.58	0.33	0.25
	H	1 or 21	10	0.41	0.21	0.20
4	L	−1 or −21	−10	0.47	0.44	0.03
	H	−10	−1 or −21	0.39	0.21	0.18

The mean over subjects was computed from the observed conditional probabilities of each subject

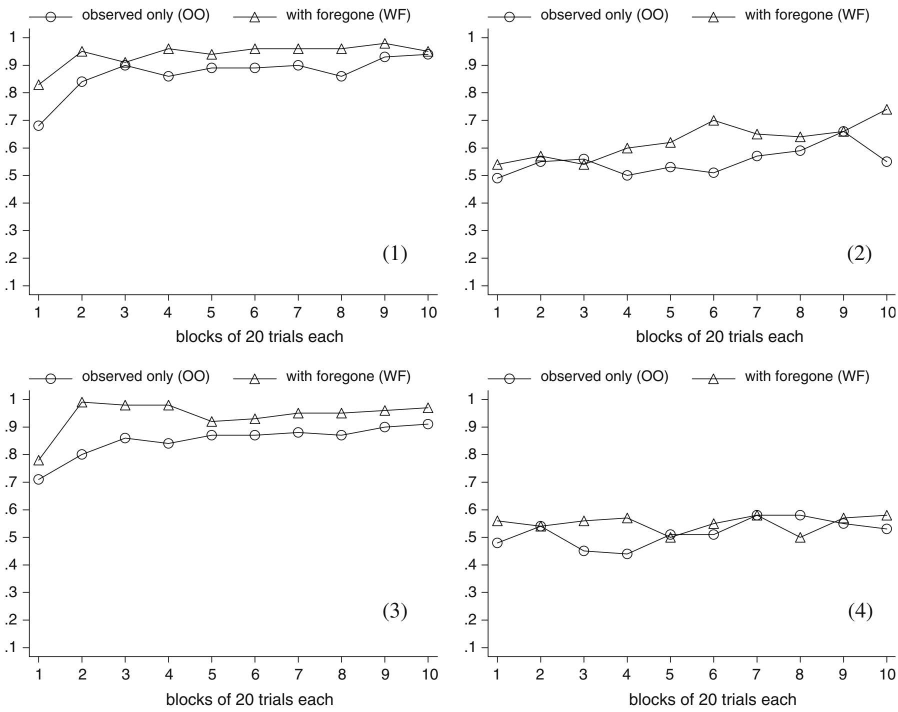


Fig. 1 Proportion of (expected) payoff maximizing choices (*H* choices) in problems 1 – 4 (10 blocks of 20 trials each)

This information reduces the risk of "getting stuck" at a sub-optimal alternative. A negative effect is predicted because in many cases (after a sequence of good outcomes from the unchosen *L* alternative) the foregone payoff information can lead to a counterproductive switch.

3 The correlation effect

Under the recency effect explanation presented above, the negative effect of foregone payoff information is expected to diminish when the payoffs obtained from the possible alternatives are positively correlated. A positive correlation implies that when an *H* alternative has high value, the *L* alternative also has high value and vice versa. Thus, there could be no situation where a low value in the *H* alternative would be displayed along with a high value in the *L* alternative. In an extreme case, in which the outcomes are perfectly correlated (and there are no sequences in which *L* has higher payoffs), only a positive effect of foregone payoff is predicted. To evaluate this prediction we examined adaptation in the following two problems:

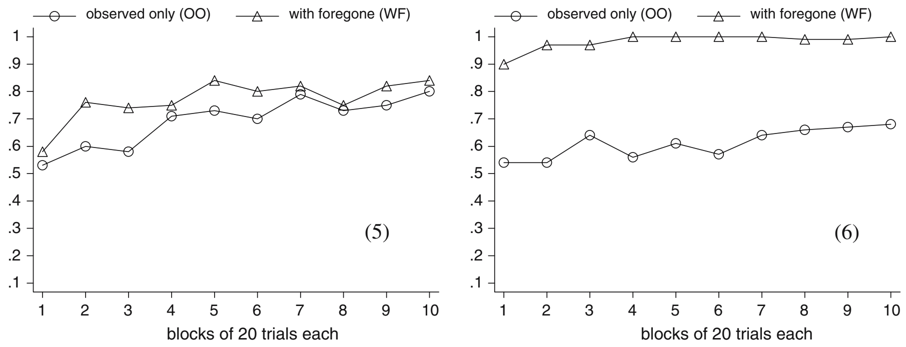


Fig. 2 Proportion of (expected) payoff maximizing choices (H choices) in problems 5 and 6 (10 blocks of 20 trials each)

Problem 5 H: Round$(h_t + u1_t)$ L: Round$(l_t + u2_t)$

Problem 6 H: Round$(h_t + u1_t)$ L: Round$(l_t + u1_t)$

where h_t is drawn from the normal distribution N(110, 20), l_t is drawn from the normal distribution N(100, 20), and $u1_t$ and $u2_t$ are drawn from the distribution (−50 with $p = 0.5$; +50 otherwise). The random variables are independently drawn in each trial. Thus, the payoffs of the two alternatives are positively correlated in problem 6, but are not correlated in problem 5. Forty Technion students (ten in each problem and information condition) participated in this study. The conversion rate was 1 Shekel per 1,000 points.

The results, presented in Fig. 2, support the predictions. The information condition had a large effect in problem 6 (correlated payoffs), but not in problem 5. Switches between H and L offset each other in problem 5 but not in problem 6, since the H alternative has (by design) a higher foregone payoff in almost all trials. Over the 200 trials the maximization rate in condition OO and WF were 0.69 (STD = 0.21) and 0.77 (STD = 0.21) in Problem 5 ($t[18] = -1.14$, n.s.), and 0.61 (STD = 0.19) and 0.98 (STD = 0.05) in Problem 6 ($t[18] = -9.88$, $p < 0.05$).[3]

[3] Diederich and Busemeyer (1999) (henceforth DB) made a similar observation regarding the correlation of outcomes. They contrasted positive with negative correlation in outcomes when foregone payoffs are provided. They find that a negative correlation leads to a lower proportion of choices of the alternative with the higher expected value. This result can be easily explained by the recency model: In the extreme negative correlation situation, low values in alternative H are almost always presented with high value in alternative L, tempting decision makers to choose from the alternative (L) with the lower expected value. This makes it more difficult to learn to choose from alternative H. Note that DB do not investigate behavior when foregone payoffs are not provided. Furthermore, they do not report the learning curves.

4 The diminishing big eyes effect

Another implication of the recency hypothesis involves the possibility that in a high variability environment with a large number of uncorrelated alternatives, information concerning foregone payoffs can impair learning. This is due to the fact that the relatively large outcomes from alternatives that have high variability can tempt decision makers to choose them. High recency contributes to the continuation of this attraction effect beyond the first presentation.

To evaluate this possibility we studied choice behavior among 100 alternatives whose payoffs were drawn from normal distributions, in which payoffs are rounded to the nearest integer (RN). Seventeen Technion students participated in each information condition.

Problem 7 100 buttons (randomly placed in a 10×10 matrix):[4]
50 H - buttons: drawn from RN(11,1) ; 50 L - buttons: drawn from RN(10,3)

The conversion rate was 1 Shekel per 100 points. Notice that in this design the highest payoff in any specific trial tends to come from alternatives with the lower mean but higher variability (the RN(10,3) distribution). Thus, overweighting of recent payoffs can lead to sub-optimal L choices. Erev and Rapoport (1998) call a similar pattern the "big eyes" effect in their study of market entry games.[5]

Figure 3 presents the observed proportion of optimal H choices (buttons with a mean of 11) in 10 blocks of 20 trials. Comparison of the two information conditions reveals an initial tendency of foregone payoffs to attract decision-makers to choose from the alternative with the high variability but low expected payoff (L).[6] However, we can see that experience decreases this tendency. The maximization rate (percentage of choices of H-buttons with a mean of 11) over the first 100 trials is 0.41 (STD = 0.23) in Condition WF and 0.53 (STD = 0.15) in Condition OO. The maximization rate over the last 100 rounds is 0.50 (STD = 0.30) in Condition WF and 0.50 (STD = 0.15) in Condition OO. The difference between the two conditions is significant in the first 100 trials ($t[32] = 1.83$, $p < 0.05$), but not in the second (and last) 100 trials ($t[32] = -.00$, n.s.).

In order to understand where the significant worse performance during the first 100 trials has its origins, we analyzed the relative frequency of all moves in both information conditions. In both conditions the relative frequency of moves is declining over time. However, subjects in the WF condition explore much more, i.e., their observed relative frequency of moves (changes in choices)

[4] See Appendix A for a screen shot of the experimental interface.

[5] They observe that when outsiders are informed about the profits of incumbents (i.e., their foregone payoffs), they are more likely to enter in the next period, reducing all entrants' profits beyond profitability.

[6] In the very first trial subjects in both information conditions select an H-button about 59% of the time. This cannot be seen in the figure, since we plot blocks of 20 trials. The selection of an H-button drops hugely during the first 20 rounds in condition WF and only slightly in condition OO.

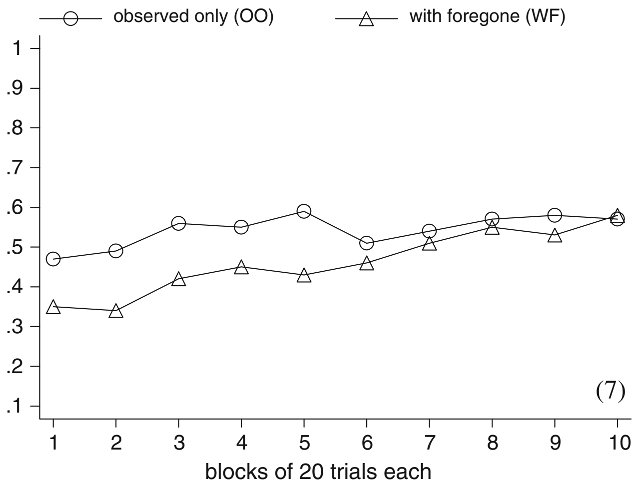


Fig. 3 Observed relative frequency of H choices in problem 7 (10 blocks of 20 trials each)

is higher than that of subjects in the OO condition. On average, they explore about 13% more during the first 100 rounds and even about 26% more in the last 100 rounds of the experiment. This high and unnecessary exploration seems to be due to the attraction of foregone payoffs. The recency effect is considered to make the effect last longer than the first presentation of the large negative payoffs in the L options.

5 The ceiling effect

The weak positive effect of foregone payoff in problems 1 and 3 (no payoff variability) can be explained with the ceiling effect. Presumably, the effect of foregone payoff due to the addition of information on the potential outcomes was insignificant because there was not much room for improvement over the high maximization rate in the OO conditions.

To analyze this suggestion we studied a choice problem including 400 alternatives (placed in a 20×20 matrix). As in problems 1 and 3, each alternative was associated with a fixed payoff that did not change throughout the 200 experimental trials. The payoff of alternative $j (j = 1, 2, \ldots, 400)$ was j. The location of the 400 alternatives in the matrix was randomly determined for each subject before the beginning of the experiment. Notice that in the current setting subjects need only one trial to find the optimum when foregone payoff information is available (Condition WF), but 400 trials are needed to find the optimum without this information (Condition OO). Moreover, since the experiment lasted only 200 trials, trying to find the optimal alternative in the OO condition is not necessarily a wise approach.

In problem 8, the conversion rate was 1 Shekel per 3,000 points. Figure 4 presents the mean of normalized payoffs in the two information conditions as a

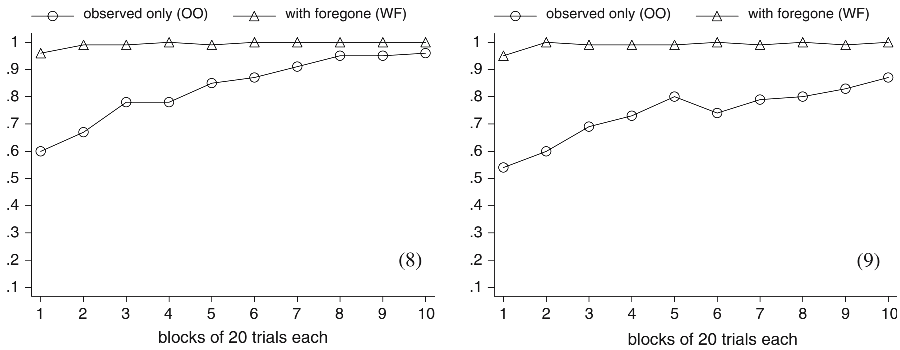


Fig. 4 Mean normalized payoffs in problems 8 and 9 (10 blocks of 20 trials each)

function of time (10 blocks of 20 trials).[7] As expected, the results reveal quick learning leading to maximized earnings by the use of foregone payoffs. The maximization rate is 0.99 (STD = 0.02) in Condition WF, and 0.83 (STD = 0.16) in Condition OO. In support of the ceiling explanation of the result of problem 1 and 3, over the 200 trials the difference is significant ($t[18] = -7.22$, $p < 0.05$).

Whereas these results are not surprising, they demonstrate the potential for the positive effect of foregone payoffs in the absence of its negative feature. Furthermore, the results are important for examining quantitative models of the task. One possible explanation of the difference between problems 7 and 8 involves the payoff magnitude. The difference between the expected payoff from optimal and random behavior was 0.005 Shekel ((11−10.5)/100) in problem 7 and 0.067 [(400−200)/3,000] in problem 8. Thus, it can be argued that the larger payoff magnitude is responsible for the positive effect of foregone payoffs. To evaluate this hypothesis we ran a replication of problem 8 with a conversion rate of 40,000 points per Shekel. With this rate the difference between random and optimal is 0.005 (as in problem 7).[8] The results reveal that the smaller payoff magnitude did not decrease the information effect. On the contrary, the difference between the two conditions in problem 9 is larger than the difference in problem 8. The maximization rate was 0.99 (STD = 0.02) in Condition WF and 0.74 (STD = 0.18) in Condition OO. This is a significant difference ($t[18] = -7.39$, $p < 0.05$).

6 Implications to models of learning

We believe that the different effects of foregone payoffs can be summarized with a "diminishing recency" assumption (see Barron and Erev 2003). Under

[7] Payoffs are normalized to lie between 0 and 1 to allow comparability to the previous problems.

[8] The experimental procedure in the replication (Problem 9) was identical to that of problem 8 with two exceptions: subjects received a 20 NIS show up fee (no show up fee was given in problem 8), and the conversion rate was 1 Shekel for 40,000 points. The show up fee was added to insure a total payoff between 20 and 22 Shekels.

this assumption the subjects initially behave as if they tend to "chase" high recent outcomes. This tendency implies a negative effect of foregone payoff information in 50% of the trials in problems 2 and 4, in about 40% of the trials in problem 5, and in most trials in problem 7; depending on the frequency that foregone payoffs in the L alternatives are attractive compared to the obtained outcomes in the H alternative. In all other problems this tendency implies a positive effect of foregone payoffs. The finding that the negative effect of foregone payoff information in problem 7 decreases with time suggests that the chasing tendency can decrease with experience.

In order to evaluate this verbal summary of the results, the following sections explore a post hoc quantification of this summary. To facilitate derivations of the implications to mainstream models of learning, the quantification examined here is a generalization of models of fictitious play (see e.g., Brown 1951; Mookherjee and Sopher 1994; Cheung and Friedman 1997; Fudenberg and Levine 1998; Nyarko and Schotter 2002; Camerer and Ho 1999; Capra et al. 2002).

Our approach is to evaluate the post hoc model based on its ability to reproduce the aggregate maximization rates using computer simulation (with a single set of parameters) in all the experimental conditions.

6.1 Fictitious play with diminishing recency

The simplest satisfactory post hoc generalization of a fictitious play model that we could develop involves three basic assumptions. The first involves a distinction between *exploration* and *exploitation* trials. As explained below, without this distinction (i.e., without exploration trials) the model is identical to a basic stochastic fictitious play model. The addition of exploration trials, allowing for faster learning and information acquisition (if necessary), and reduces the tendency of basic models to "get stuck" in sub-optimal action when foregone payoff information is not available. Thus, this addition approximates the prescription of the optimal rule for Bandit problems (see Berry and Fristedt 1985). Like the optimal prescription in Bandit problems the tendency to explore is a function of the number of trials to be played.

Diminishing exploration The probability that a particular trial is an exploration trial decreases with time. A linear function is assumed.

$$P(t \text{ is an exploration trial}) = \frac{r\kappa}{(t + r\kappa)}, \tag{1}$$

where r is the expected number of trails in the experiment (200 in the current study), and $\kappa > 0$ is an "exploration" parameter.

Contingent weighting The second assumption allows for the possibility of higher weighting for the outcomes observed in exploration trails. This assumption is added to capture the intuition that even a few exploration trials can be important. The adjusted value of alternative j at trial $t+1$ is:

$$q_{j,t+1} = \left(1 - w_{jt}\right) q_{jt} + w_{jt} \cdot x_{jt}, \tag{2}$$

where x_{jt} is the payoff of j at t, and w_{jt} determines the weight of this payoff. The initial value, q_{j1}, is the expected payoff from random choice. The weight w_{jt} is assumed to depend on whether j's payoff was observed, and whether trial t was an "exploration" or "exploitation" trial:

$$w_j t = \begin{cases} 0 & \text{if } j's \text{ payoff was not observed} \\ \delta\theta & \text{if } j's \text{ payoff was observed and } t \text{ is an exploitation trial,} \\ \theta & \text{if } j's \text{ payoff was observed and } t \text{ is an exploration trial} \end{cases} \tag{3}$$

where $0 < \theta < 1$ is the adjustment speed in exploration trials, and $\theta\delta(0 < \delta \leq 1)$ is the (possibly slower) adjustment speed in exploitation trials.

Contingent choice rule The final assumption implies a stochastic response rule. When the foregone payoffs are not known, the decision maker is assumed to randomly select one of the alternatives in exploration trials. In all other cases (known foregone payoffs and/or exploitation trials) the following choice rule is applies:

$$p_k(t) = \frac{\exp(\lambda q_k(t))}{\sum_{j \in J} \exp(\lambda q_j(t))}. \tag{4}$$

6.1.1 Descriptive value

The second column of Figs. 5, 6 and 7 presents the predictions of the generalized FP model. The predictions were derived using computer simulations, in which 100 virtual agents that behave according to the model's assumption participate in a replication of each of the experimental condition. To estimate the four parameters we use a grid search procedure with mean squared deviation criteria (see Selten 1998), i.e., we run simulations with different set of parameters to find the set of parameters that minimize the MSD between the observed and simulated curves. The best fit (MSD = 0.012) was obtained with the parameters $\kappa = 0.15, \lambda = 200, \theta = 0.17$ and $\delta = 0.25$. As we can see from Figs. 5, 6 and 7, using these parameters the model reproduces the main experimental results.

6.1.2 The contribution of the post hoc addition

With the constraint $\kappa = 0$ (no exploration) the model presented above is reduced to a weighted FP model. This constraint increases the MSD score by a factor of four to 0.045 (and the estimated parameters are $\lambda = 100, \delta\theta = 0.12$, notice that with $\kappa = 0$ only the product of δ and θ is relevant). The bad fit of the model without exploration appears to be a result of its failure to capture, with the same set of parameters, the fast learning in problems 1, 3, 8 and 9 and the results in problems 4 and 7. The parameters that imply fast learning in 1, 3, 8 and 9, imply maximization rates close to zero in problem 2 without foregone payoff information and in problem 7 with foregone payoff information.

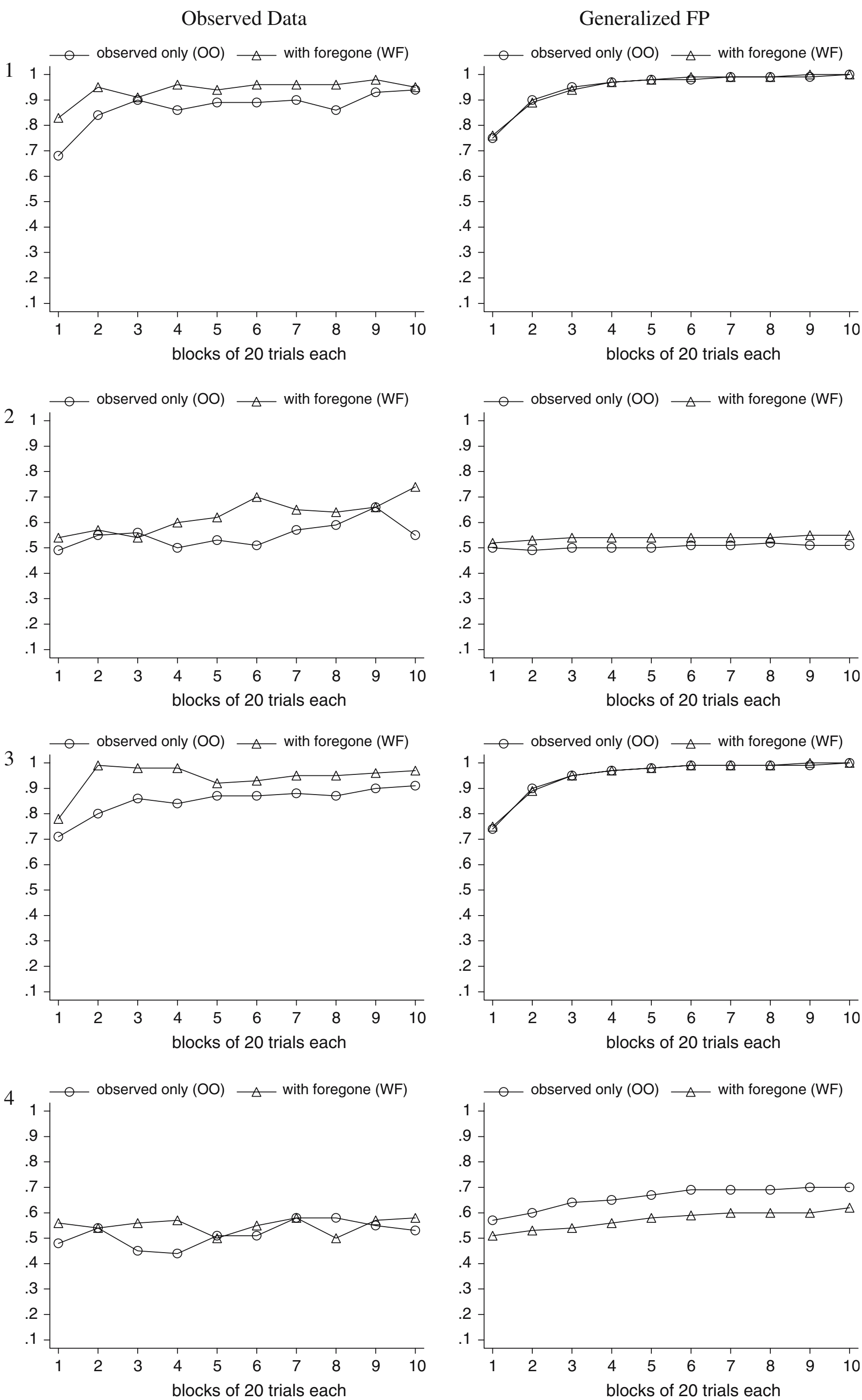


Fig. 5 Observed and simulated behavior in problems 1 through 4

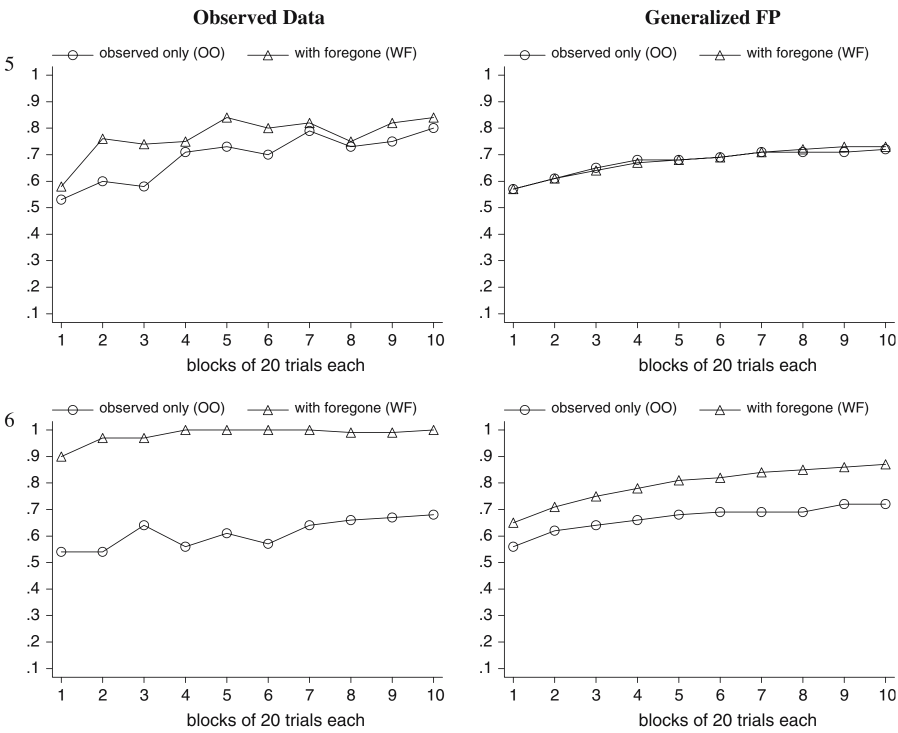


Fig. 6 Observed and simulated behavior in problems 5 through 6

The constraint $\delta = 1$ implies the same recency effect in exploitation and exploration trials (and as such over time). This constraint increases the MSD score to 0.015 and the estimated parameters are $\kappa = 40, \lambda = 100, \theta = 0.12$. The payoff sensitivity parameter λ is less important. Indeed the constraint $\lambda \to \infty$ (deterministic choice of the alternative with the highest adjusted value in exploitation trials) does not impair the fit. Notice that with this constraint the current model is a variant of Sarin and Vahid's deterministic Payoff Assessment model (Sarin and Vahid 1999, 2001).[9]

6.2 The money-pump critique, and reinforcement learning among strategies

An important shortcoming of the model presented above emerges in an attempt to derive its prediction to situations in which greedy "mechanism designers" try to pump their mechanism's users' money. One example involves a casino designer who offers a choice between 100 options. Assume that 99 of the options

[9] In addition to these basic fictitious play models we also examined the more complex experience weighted attraction (EWA) proposed by Camerer and Ho (1999). The results reveal that in the current setting EWA is not more useful than the basic models (see Haruvy and Erev 2001). Indeed, its estimated parameters would imply fictitious play.

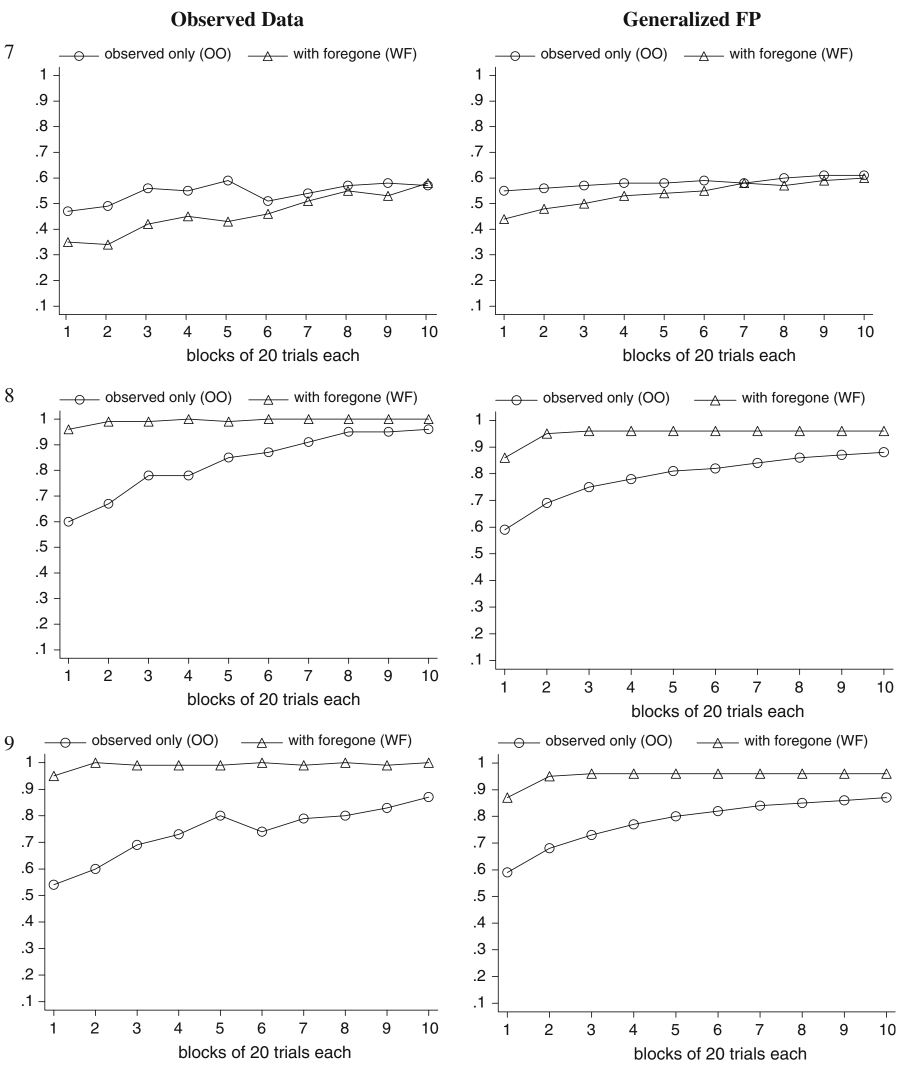


Fig. 7 Observed and simulated behavior in problems 7 through 9

are independent draws of a gamble that pays 200 with $p = 0.01$, and -3 otherwise. Option 100 is the status quo (0 with certainty). After reading the previous section, the designer elects to condition the availability of foregone payoffs on their realizations. In the designed casino, foregone payoffs become publicly known only if they are positive.

It is easy to see that individuals who behave according to the current model (or other fictitious play models) will find gambling attractive. Diminishing exploration will not help in this case: The model predicts 100% gambling rates in the long term. Similar thought experiments led Erev and Roth (1998, 1999) to propose that a reinforcement-learning model (that focuses on obtained payoffs) is likely to be more robust than fictitious play models. However, basic reinforcement learning model (e.g., Roth and Erev 1995) cannot account for the clear effects of foregone payoff summarized here. To address this problem Erev and Roth (1999) proposed the idea of reinforcement learning among cognitive strategies. When fictitious play is one of the cognitive strategies, the agents can learn to select it when it is reinforcing, and learn to stop using it when it leads to ineffective outcomes.

7 Discussion

The current paper demonstrates that the effect of foregone payoff information on human adaptation "changes with the wind." Very different effects were observed in different environments. These effects can be divided into positive (maximization rate enhancing) effects and negative (maximization rate reducing) effects.

Positive effects of the availability of foregone payoff information are the result of increased information. Clear positive effects were observed in three problems: (1) in a low variance environment with a large number of possible alternatives (problems 8, 9), and (2) in an environment where the payoffs to the alternatives are positively correlated (problem 6).

However, the provision of foregone payoffs can also have a negative effect due to momentary attraction to foregone payoffs (problem 7). This initial attraction can increase the subjective value of a highly variable but risky alternative that from time to time has high payoffs. If foregone payoffs are not provided, decision makers are not exposed to such high value outcomes if they have not chosen them. Our data suggest that this negative effect decreases with experience. Finally, in five problems (problems 1–5) the maximization rate was found to be insensitive to the availability of information concerning foregone payoffs.

Whereas these patterns cannot be captured with existing learning models, it is possible to construct post hoc models of learning that summarizes all the patterns with one set of parameters. The current paper presents one example, a model of fictitious play with diminishing exploration, and discusses the feasibility of other examples that assume reinforcement learning among cognitive strategies.

7.1 Potential implications

In order to highlight the potential implications of the behavioral regularities summarized above we chose to conclude with a discussion of the some examples. As mentioned earlier, a designer of new economic environments can often control the feedback received by agents. In particular, designers can try to limit the feedback to information concerning obtained payoffs only, or allow information concerning foregone payoffs.[10] One example is provided by lotteries. Problem 7 of our paper suggests that the availability of foregone payoff information is expected to facilitate maladaptive gambling. Operators of lotteries seem to agree. For example, they spend a lot of money on presenting the winning number on TV (rather than providing a cheaper and more effective free phone line). Another consistent phenomenon concerns the expected value (EV) from different gambling games. In Israel, for example, the EV of the main state lottery with natural foregone payoff information (like the numbers games, in which gamblers can pick their numbers) is 51% of investment and the EV of lotteries without natural foregone payoffs (when each player can only select among tickets and only knows whether he has won or not) is 61% of the investment.

Finally, sensitivity to recent foregone payoffs implies that investors would chase successful funds. Barber et al. (2004) find support for this prediction. In a study of American households' investment behavior they find that mutual funds purchases mainly occur in the top quintile of past annual returns. In a similar context, Benartzi and Thaler (1995) explained the equity premium puzzle by myopic loss aversion (MLA). They demonstrated that rejecting single gambles but accepting a sequence of such gambles is consistent with MLA. If returns are evaluated over a longer period of time, multiple gambles become more attractive due to the lower probability that a loss will be experienced. Therefore the period over which individuals evaluate financial outcomes influences their investment in risky assets.[11] Our paper suggests a complimentary explanation, namely that a longer evaluation period can be a mechanism to avoid the chasing of attractive (but potentially misleading and short-termed) foregone payoffs.

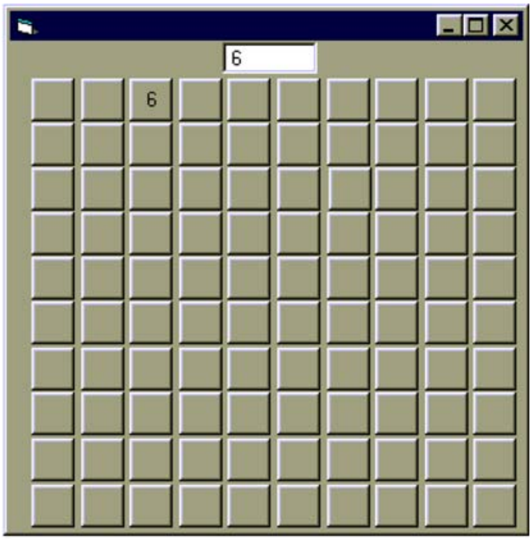
In summary, the current analysis clearly demonstrates that foregone payoff information affects choice behavior in significant ways, depending on the environment. Future theoretical work should be more perceptive to this influence. Our results further suggest that although the effect of foregone payoff information appears to change with the wind, it can be predicted with simple models that quantify robust principles. Interestingly, the current attempt to capture the observed behavioral pattern led to a generalization of the fictitious play model that approximates the predictions of the optimal choice in bandit problems. This observation can be used to refine models of human adaptation, and can help shed light on interesting mechanism design problems.

[10] Notice that the design problems we consider differ from the problems addressed in mainstream mechanism design research. Mainstream research (e.g. Myerson 1989, chapter 23 of Mas-Colell et al. 1995) tends to address situations, in which agents are expected to behave rationally and the behavior of rational agents can be unambiguously derived. In the problems we consider, agents might not behave rationally, and as we noted above, the derivation of the rational behavior is not necessarily obvious.

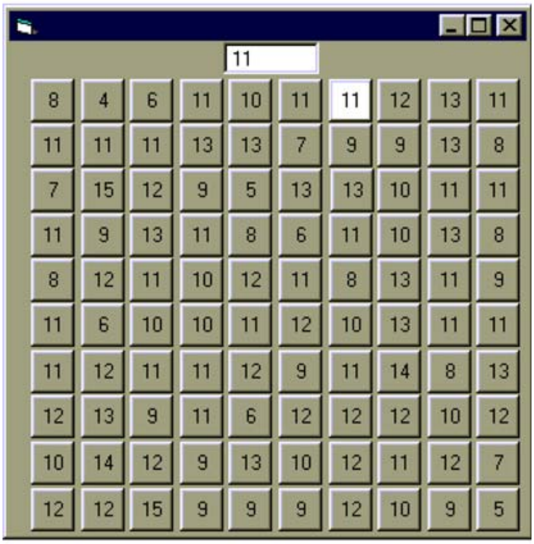
[11] A similar observation was made by Gneezy and Potters (1997).

Appendix

B1: Computer setup of problem 7



Observed-Only Condition



With Foregone Condition

References

Barber B, Odean T, Zheng L (2004) Out of sight, out of mind – the effect of expenses on mutual fund flows. J Business (in press)

Barron G, Erev I (2003) Small feedback based decisions and their limited correspondence to description based decisions J Behav Decis Making 16:215–233

Benartzi S, Thaler R (1995) Myopic loss aversion and equity premium puzzle. Q J Econ CX:73 – 92

Berry D, Fristedt B (1985) Bandit problems: sequential allocation of experiments. Chapman and Hall, London/New York

Brown G (1951) Iterated solution of games by fictitious play. In: Koopmanns TC (ed) Activity analysis of production and allocation. Wiley, New York, pp. 374–376

Camerer C, Ho T (1999) Experience-weighted attraction learning in games: a unifying approach Econometrica 67:827–874

Capra M, Goeree JK, Gomez R, Holt C (2002) Learning and noisy equilibrium behavior in an experimental study of imperfect price competition. Int Econ Rev 43(3):613–636

Cheung YW, Friedman D (1997) Individual learning in normal form games: some laboratory results. Games Econ Behav 19:46–76

Diederich A, Busemeyer JR (1999) Conflict and the stochastic dominance principle of decision making. Psychol Sci 10: 353–359

Erev I, Rapoport A (1998) Coordination, 'magic,' and reinforcement learning in a market entry game. Games Econ Behav 23(2):146–175

Erev I, Roth A (1998) Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. Am Econ Rev 88(4):848–881

Erev I, Roth A (1999) On the role of reinforcement learning. In: Budescu DV, Erev I, Zwick R (eds) Games and human behavior: essays in honor of Amnon Rapoport. LEA, Hillsdale, pp. 53–77

Fudenberg D, Levine DK (1998) The theory of learning in games. MIT Press Cambridge

Gneezy U, Potters J (1997) Risk taking and evaluation periods. Q J Econ CXII:631–645

Haruvy E, Erev I (2001) Interpreting parameters in learning models In: Zwick R, Rapoport A (eds) Experimental business research. Kluwer, Dordrecht, pp. 285–300

Heller D, Sarin R (2002) Adaptive learning with indirect payoff information. mimeo

Mas-Colell A, Whinston M, Green J (1995) Microeconomic theory. Oxford University Press, New York

Mookherjee D, Sopher B (1994) Learning behavior in an experimental matching pennies game. Games Econ Behav 7:62–91

Myerson R (1989) Mechanism design. In: Eatwell J, Milgate M, Newman P. (eds) Allocation, information and markets. W.W. Norton, New York, pp 191–206

Nyarko Y, Schotter A (2002) An experimental investigation of belief learning with real beliefs. Econometrica 70(3):971–1005

Robinson J (1951) An iterative method of solving a game. Ann Math 54:296–301

Roth A, Erev I (1995) Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. Games Econ Behav 8(1):164–212

Sarin R, Vahid F (1999) Payoff assessment without probabilities: a simple dynamic model of choice. Games Econ Behav 28:294–309

Sarin R, Vahid F (2001) Predicting how people play games: a simple dynamic model of choice. Games Econ Behav 34:104–122

Selten R (1998) Axiomatic characterization of the quadratic scoring rule. Exp Econ 1:43–62