A Laboratory Study of Bayesian Updating
in Small Feedback-Based Decision Problems

[1]Takemi Fujikawa and [2]Sobei H. Oda
[1]School of Economics and Finance, University of Western Sydney, Australia
[2]Faculty of Economics, Kyoto Sangyo University, Japan

Abstract: This study explores small feedback-based decision problems experimentally. Conducted were the experiments in which the decision-maker's payoff distribution was limited to either favorable distribution or unfavorable distribution. The first remarkable observation revealed complexity/loss aversion in the experiment. The second observation included the law of small numbers. Deviations from maximization were also observed. Finally, we investigated the imperfect Bayesian decision-makers observed in the experiment by exploring to what extent the decision-makers could update subjective Bayesian probability and rely on it in making decisions.

Key words: Sequential search, Bayesian updating, small feedback-based decision problems, experiment, JEL Classification: C91, D81, D83

INTRODUCTION

This study conducts search experiments on Small Feedback-based Decision problems (SFD). SFD are defined as consequential decision problems but each single choice is not very important because the options available to the Decision-Maker (DM) have similar expected values that may be quite small, so that little time and effort is typically invested in these problems[1]. The DM in SFD is supposed to make his decision many times without evaluating carefully the possible outcomes.

This research carries out extensive experimental exploration of the process of Bayesian updating with SFD. There has been some literature about search experiments on SFD[1,2], none of this literature has, however, focused upon the process of Bayesian updating. This study conducts search experiments focusing upon the DM's sequential search process of Bayesian updating on SFD.

The current experiments were conducted with the repetition of 400 rounds, while many previous experiments[3] focused upon one-shot description-based decisions. The reason of conducting repeated-play conditions is that economics experiments typically use stationary replication, where the same task is repeated over and over, with fresh endowments in each period. Data from the last few periods of the experiments are typically used to draw conclusions about the equilibrium behavior outside the laboratory[4].

Present results exhibit the DMs' remarkable tendencies. The first remarkable observation reveals the DMs' complexity/loss aversion in the experiment. The second is that the DMs behave as if the law of small numbers is revealed. Deviations from maximization (low maximization) are also observed. The third observation is that the DMs behave as if they are imperfect Bayesians.

Bayesian updating: The standard principles adopted in economics to model probability judgment under uncertainty are concepts of Bayesian updating. Bayesian updating helps us concern the manner in which the DM processes new information and update his beliefs.

Consider a game in which the following two equally likely states of the world are available to the DM, a priori relatively high state, State A and a priori relatively low state, State B. Let α, $\beta > 0$, p_1, $p_1 \in [0, 1]$, $\alpha p_1 > \beta$ and $\alpha p_2 > \beta$. In State A, two bingo cages are available: cage H from which a ball numbered α is drawn with probability p_1; cage L a ball numbered β with certainty. In State B, two bingo cages are available: cage H from which a ball numbered α is drawn with probability p_2; cage L a ball numbered β with certainty.

State A. Choose between:
H: α points with probability p_1; 0 otherwise
L: β points with certainty.

State B. Choose between:
H: α points with probability p_2; 0 otherwise
L: β points with certainty.

At the beginning of the game, the DM is presented with the two equally likely states of the world introduced above and its payoff structure. The DM is asked to choose for 400 times one of the two cages,

Corresponding Author: Takemi Fujikawa, School of Economics and Finance, University of Western Sydney,
Email: t.fujikawa@uws.edu.au

cage H or cage L, from which one ball is drawn at a time. It is undisclosed which of the two states of the world is an actual state throughout the game, however, disclosed that the same state of the world yields draws over 400 trials. Hence, the DM will be expected to discover which of the two states of the world be realized actually.

We explore an analysis in this study the assumption that the rational DM should make his decision to maximize his expected payoff (utility) under uncertainty This assumption asserts that the DM is willing to keep choosing H (L) after he has appeared to an actual state to be State A (B).

At period t, the DM's updated probability of recognizing an actual state of the world in the process of Bayesian updating, facing the outcome, xt, is given as:

$$P_t(StateA \mid x_t) = \begin{cases} \frac{P(\alpha|StateA)P_t(StateA)}{P(\alpha|StateA)P_t(StateA)+P(\alpha|StateB)P_t(StateB)}, & \text{if } x_t = \alpha \\ \frac{P(0|StateA)P_t(StateA)}{P(0|StateA)P_t(StateA)+P(0|StateB)P_t(StateB)}, & \text{if } x_t = 0 \\ P_{t-1}(StateA), & \text{if } x_t = \beta \end{cases}$$

From the tenets of Bayesian updating and a rationality assumption, we propose the following important hypothesizes on the DM's behavior. One is that the DM should choose an alternative H whenever $P_t(StateA/x_t) > 0.5$ at period t, implying that State A is more likely to be an actual state for the DM. The second hypothesis is that the DM should choose an alternative L whenever $P_t(StateA/x_t) < 0.5$ at period t, implying that State B is more likely to be an actual state for the DM.

MATERIALS AND METHODS

Two economics experiments, Experiment 1 and 2 were conducted at Kyoto Sangyo University Economic Experiment Laboratory. Thirty-three undergraduates at Kyoto Sangyo University participated in both Experiment 1 and 2 in order. Both Experiment 1 and 2 were conducted under the condition that the subjects were informed of the exact number of rounds and sessions to be performed. The subjects received monetary payoffs according to the exchange rate: 1 point= 0.6 Yen (0.5 US cent).

Both in Experiment 1 and 2, the subjects were asked to join four sessions, Session 1, 2, 3 and 4, each of which was consisted of 400 rounds (100 rounds only in Session 1) under the condition that the subjects were presented with two equally likely states of the world at the beginning of each session, a priori relatively high state (good news) and a priori relatively low state (Bad news). The subjects were undisclosed an actual state of the world during each session, however, were disclosed that the same state of the world was yielding draws across one session. Hence, the subjects were expected to discover which of the two states of the world was actually generating each draw in each session.
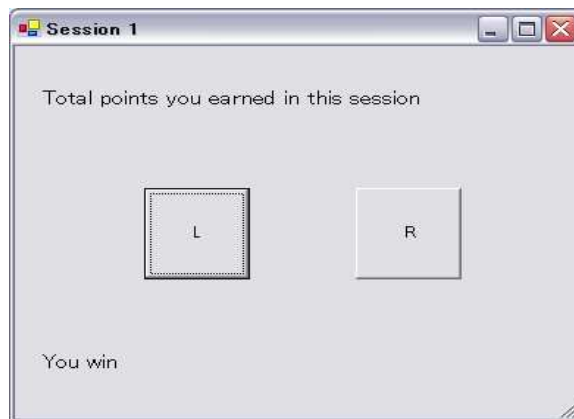

Fig. 1: Computerized money machine

Throughout both Experiment 1 and 2, the subjects were instructed to operate a computerized money machine shown in Fig. 1. The subjects' basic task at each round was a binary choice between L and R for 400 times in each session. The payoff structure of the two buttons is introduced in the following section. Among both experiments, the money machine provided the subjects with binary types of feedback immediately following each choice: the payoff for the button chosen, that appeared on the screen for the duration of one second and an update of an accumulating payoff counter, which was constantly displayed.

Experiment 1: In Experiment 1, the subjects were provided with two equally likely states of the world: State A (good news) and State B (bad news), however they were undisclosed that State A was a dummy state and therefore State B was an actual state for all of the four sessions. Let (V, p) be an alternative that yields a payoff of V points with probability p and zero otherwise:

Session 1: State A: Choose between L: (6, 1) and R: (5, 1).
State B: Choose between L: (4, 1) and R: (3, 1).
Session 2: State A: Choose between L: (4, 0.9) and R: (3, 1).
State B: Choose between L: (4, 0.8) and R: (3, 1).
Session 3: State A: Choose between L: (4, 0.3) and R: (3, 0.25).
State B: Choose between L: (4, 0.2) and R: (3, 0.25).
Session 4: State A: Choose between L: (32, 0.2) and R: (3, 1).
State B: Choose between L: (32, 0.1) and R: (3, 1).

Experiment 2: The setting for Experiment 2 is the same as Experiment 1 with the exception that State A

was an actual state for all of the four sessions and it was undisclosed to the subjects.

Session 1: State A: Choose between L: (4, 1) and R: (3, 1).
State B: Choose between L: (2, 1) and R: (1, 1).
Session 2: State A: Choose between L: (4, 0.8) and R: (3, 1).
State B: Choose between L: (4, 0.7) and R: (3, 1).
Session 3: State A: Choose between L: (4, 0.2) and R: (3, 0.25).
State B: Choose between L: (4, 0.1) and R: (3, 0.25).
Session 4: State A: Choose between L: (32, 0.1) and R: (3, 1).
State B: Choose between L: (32, 0.05) and R: (3, 1).

RESULTS AND DISCUSSION

Table 1 shows the mean proportion of L choices (choice L) throughout 400 rounds in each session in Experiment 1 and 2. Figure 2-4 shows the choice L in blocks of 50 trials to facilitate an efficient summary of the large set of the data. On the one hand, among both experiments, we see that the reversed certainty effect was observed in Session 2 since choice L were more than 0.5. On the other hand, it is found that the choice L in Experiment 1 was significantly larger than that in Experiment 2 for all of the four sessions. The corresponding p-values are 0.491, 0.000, 0.319, 0.460 for Session 1, 2, 3, 4 respectively.

L or R starts choice: We find pronounced tendency that there were fewer L start choices in Experiment 2 than in Experiment 1, in phases where our subjects made their starting choices. The mean proportion of L start choices were 0.68, 0.94 and 0.65 in Session 2, 3 and 4 in Experiment 1 respectively, while 0.39, 0.67 and 0.58 in Experiment 2. The difference between the two experiments is significant (P (Z ≤ z) =0. 069).

Regarding Session 2 in Experiment 2, there exists remarkable tendency that R was chosen often as starting choice as though complexity/loss aversion was exhibited in the first trial in spite of the following two facts. One is that in being made the first draw, both L and R offers the same expected payoff if there is only one draw. If there is only one draw, the expected payoff (utility) of the two alternatives is the same:

$$EU(L) = \frac{1}{2} \times \{4 \times 0.8\} + \frac{1}{2} \times \{4 \times 0.7\} = 3, \quad EU(R) = 3$$

The second fact is that observing the outcome of the first R draw does not resolve uncertainty regarding the state of the world (good or bad news) as the first L

draw may do.

Regarding Session 3 in Experiment 2, there exists a substantial tendency that L was chosen often as starting choice although L offered less expected payoff than R did if there was only one draw. If there is only one draw, the expected payoff (utility) of L is lower than R:

$$EU(L) = \frac{1}{2} \times \{4 \times 0.2\} + \frac{1}{2} \times \{4 \times 0.1\} = 0.6, \quad EU(R) = 0.75.$$

This trend is a mirror image of complexity/loss aversion as in Session 2 in Experiment 2. In addition, as regards both Experiment 1 and 2, the proportion of L start choices in Session 3 was the highest of all the sessions.

Table 1: The mean proportion of L choices throughout 400 rounds

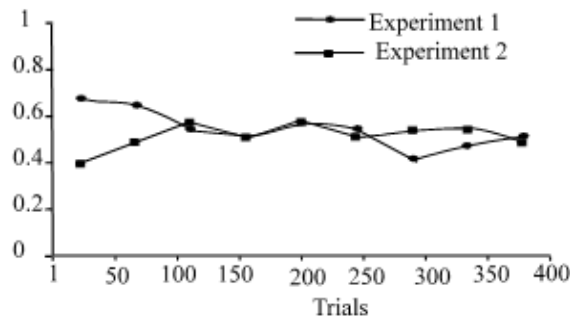|	Session 1	Session 2	Session 3	Session 4
Experiment 1	0.944	0.56	0.76	0.52
Experiment 2	0.94	0.54	0.5	0.46

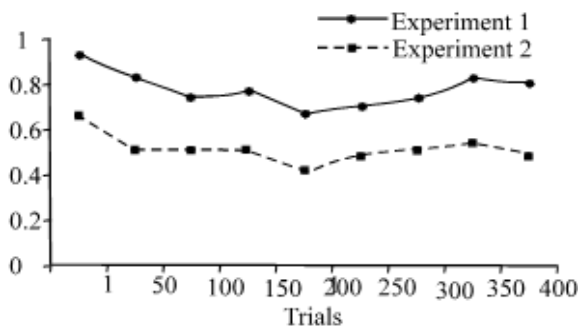Fig. 2: Choice L in blocks of 50 trials in session 2

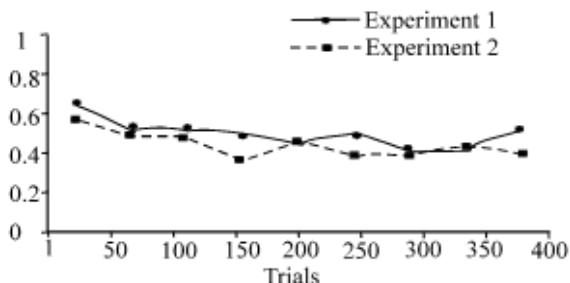Fig. 3: Choice L in blocks of 50 trials in session 3

Fig. 4: Choice L in blocks of 50 trials in session 4

Regarding Session 4 in Experiment 2, there exists tendency that more subjects, on average, started with L in Session 4 in spite of the fact that L offered less expected payoff than R did if there was only one draw:

$$EU(L) = \frac{1}{2} \times \{32 \times 0.1\} + \frac{1}{2} \times \{32 \times 0.05\} = 2.4, \quad EU(R) = 3.$$

One explanation of this tendency is that the subjects were likely to *overweight* small probabilities at the beginning of Session 4. The alternative R, however, was chosen often gradually as the subjects obtained binary types of feedback repeatedly throughout 400 rounds either in the process of "adaptive learning" or on account of the effect of the expectation of playing gambles repeatedly.

It is particularly interesting to focus on the DM's process of Bayesian updating after an initial draw in Experiment 2. One implies that after having a successful outcome in the first round, an outcome of 4 in Session 2 and 3, or 32 in Session 4, the Bayesian maximizes expected utility DM should stay with L; after having an unsuccessful outcome in the first round, an outcome of 0 in Session 2, 3 and 4, the DM should switch to R. The current results show that after having a successful outcome in Session 2, an initial outcome of 4, most of the subjects (91%) updated well and preferred to stay with L as the above hypothesis suggests. Remarkably, all subjects who had received an initial draw of 32 in Session 4 preferred to stay with L. On the other hand, all subjects in Session 2 and all but quite fewer of the subjects in Session 4 (94%) updated mistakenly and kept staying with L after receiving the unsuccessful outcome, an initial draw of 0.

The law of small numbers: The law of small numbers was observed in both Experiment 1 and 2. The law of small numbers posits that the DM will gather too little data and over generalize from small samples to distributions[6]. Assuming that the rational DM should behave to maximize his expected payoff under uncertainty, the DM's over generalization of a payoff distribution may sometimes lead him to behave irrationally. In economic applications, each DM will search too little and learn too quickly, compared to models of optimal sampling and inference[7]. One would insist that too little search leads the DM to learn mistakenly and mistaken learning induces the DM to behave irrationally.

The current results indicate that the DM chose the alternative too little and learn mistakenly too quickly. Table 1, shows that the subjects in Session 4 in Experiment 2, on average, chose L only 184 out of 400 times. One possible explanation of this is that the subjects might try L too little (only 184 times) and learn mistakenly too quickly that L had less expected payoff than R. Mistaken learning is likely to induce the subjects to choose R many times.

Over weighting and underweight small probabilities: There has been some literature on salient properties of over weighting and under weighting of rare probabilities in both one-shot description-based decisions and (repeated) SFD. Firstly, Kahneman and Tversky[3] found with questionnaire-based experiments that the average DMs in one-shot description-based decisions behaved over weighting small probabilities. Most of the subjects over weighting generally low probabilities preferred the gamble (5000, with p = 0.001; 0 otherwise) over a sure payoff with the same expected payoff. Secondly, Barron and Erev[2] found that the average DMs in SFD behaved as if they under weighted small probabilities and most DMs preferred the risk-less gamble, which yielded 3 with certainty, over the gamble (32, with p = 0.1; 0 otherwise).

Low maximization rates were observed in our experiments except Session 1. This observation is the reverse of the one in the description-based decision experiment conducted by Kahneman and Tversky[3]. It is insisted that the effect of the expectation of playing gambles repeatedly leads to the low maximization rates observed in the current experiments. Note that Kahneman and Tversky's subjects were asked to perform choice problem only once with exact prior information on payoff structure and paid hypothetical payoffs; Barron and Erev's subjects were asked to perform choice problem 400 times repeatedly without any prior information on payoff structure and paid monetary payoffs. Our results show a similar trend to Barron and Erev's results indicating under weighting of rare events in SFD, contrary to one-shot description-based decisions. It is straightforward for the subjects in Session 4 in Experiment 2 to choose L often, revealing deviations from expected payoff (utility) maximization in SFD.

Imperfect Bayesians: Some of the subjects appeared to be imperfect Bayesians. This section explores to what extent the subject in Experiment 2 can update his Bayesian updated subjective probability of recognizing an actual state of the world in Experiment 2 (updated P) and rely on the DM's updated P in making his decisions. This exploration can be done technically by investigating a correlation between the updated P and choice L. We represent in Fig. 4 the aggregated subjects' updated P and choice L in blocks of 50 trials.

The current results reveal that the subjects' mean updated P remained more than 0.5 after T = 1 in Session 2, while after T = 12 in Session 4. One set of implications is concerned with that the maximum of 400 trials should be sufficient for the DM for judging an actual state of the world in Experiment 2 correctly. That is, the subject could update his posterior information that each draw following would be coming from State A with probability of more than 0.5 after choosing L at T = 1 and T = 12 in Session 2 and 4 respectively.

The current results also reveal that the subjects, on average, never kept choosing L after T = 1 and T = 12 in Session 2 and 4 respectively in spite of the fact that the subjects' mean updated P remained more than 0.5 after those periods. One implies with this result that the subjects appeared to be imperfect Bayesians and less-than-fully-rational DMs on the ground of being unconditional upon their updated P in forming beliefs over a state of the world. A rationality assumption asserts that the perfect Bayesian rational DM should keep choosing L whenever his updated P are more than 0.5 in order to maximize expected payoff (utility).

Methodologies: One insists that a SFD experiment should be conducted with the condition that the choices and payoffs of others can be observed to each DM. In spite of the above, the current experiments were actually conducted in the setting that each DM was informed of no information as to others' choices and payoffs. This is likely setting on the ground that in many routine-learning models, knowing others' choices and payoffs is inessential since the DM is assumed to simply choose strategies that yielded high payoffs in the past[6].

Another insists that a SFD experiment should be conducted under the condition that each DM is questioned in each trial which of the two states of the world is the actual one to be realized. This should be to the point at a rough glance but we have considered it inappropriate settings for the current experiment due to the following reasons. Firstly, one considers it unreasonable setting that the DM is asked to answer repeated questions, which are not experimenter's primary concerns and may affect DM's decision making either directly or indirectly. Recall that the primary concern of our SFD experiment is not to ask which of the two states of the world the DM should consider to make a decision in each trial, but to observe what alternative the DM chooses. Secondly, asking the DM either State A or B many times (for 1300 times in each experiment) will take the DM much time and effort and induce careful evaluation of the possible options in the DM's decisions. Although careful evaluation is needed in big description-based decision experiment, we should avoid such careful evaluation in SFD experiment. Lastly, the main concern in this study is that repeated questions in each trial are likely to influence the DM's adaptive learning for making his optimal decision.

CONCLUSION

We have examined decision making on SFD in a laboratory experiment. The DM's search propensity has been explored in the context of Bayesian updating and some simple econometric methods have been employed in this study.

Further research on a search under uncertainty would clarify the following two issues. The first issue concerns to what extent the DM relies upon updated P in making choices. The second issue concerns to what extent the DM makes use of a naive heuristic in making choices.

To the best of our knowledge, there has been no literature, which aims at reviewing econometric studies on the DM's individual search behavior in SFD that use data from national economies. Yet it is straightforward to use search and choice models as maintained hypotheses for conducting econometric estimation. Hence it is hoped that further research on this type of decision making in SFD would clarify the empirical validity of search theory itself.

REFERENCES

1. Barron, G. and I. Erev, 2003. Small feedback-based decisions and their limited correspondence to description-based decisions. J. Behavioral Decision Making, 16: 215-233
2. Fujikawa, T. and S. H. Oda, 2003. Decision making under uncertainty and ambiguity. Paper presented at the The 7th Exp. Economy. Conf. Japan, Kyoto, Japan.
3. Kahneman, D. and A. Tversky, 1979. Prospect theory: An analysis of decision under risk. Econometrica, 47: 23-53.
5. Becker, G. S. and K. M. Murphy, 1988. A theory of rational addiction. J. Political Economy, 96: 675-700.
4. Camerer, C. F., G. Loewenstein and M. Rabin, (Eds.), 2004. Advances in behavioral economics. New Jersey: Princeton Univ. Press.
6. Tversky, A. and D. Kahneman, 1971. The belief in the law of small numbers. Psychol. l Bull., 76: 105-110.
7. Kagel, J. H. and A. E. Roth, (Eds.), 1995. The handbook of experimental economics. New York: Princeton Univ. Press.