Random Walk Perception and Information Acquisition in an Interactive Prediction Task Experiment

Hiroshi Izawa[*]
Mardyla Greg[†]

Abstract

We investigate experimentally individual random walk perception biases and the existence of decision clustering in a simple interactive prediction task. Our design presents a series of sequential choice problems in which the subjects are asked to forecast the subsequent outcome of a discrete binary random process. The data is generated in such a way that observation of other participants' cumulated choices makes it possible to obtain a more precise estimate of the probability distribution governing the outcomes. This setup mimics a stock market in which observing the order book provides information about possible existence and direction of a trend in prices.

We are mostly interested in the timing of subjects' decisions – a binary choice of a single purchase or sale of a security within a finite time sequence based on acquired information. Our data points to some compelling insights into rationality of Bayesian updating. Majority of our subjects display a type of irrational impatience: in tasks where they should optimally learn as much information as possible and wait until the last period to decide, they "pull the trigger too fast", incurring excessive decision costs, even when allowed to freely observe others' choices. This finding contrasts with a setting where more explicit delay costs are incorporated. Additionally, we find no apparent evidence of decision clustering or endogenous "herding" when others' actions are observable.

1. Introduction

The objective of this research is to investigate experimentally the existence of and the extent of decision clustering in a simple binary interactive discrete prediction task. In particular, we are interested in examining the following issues:

- Do decision-makers perceive uncertainty when there is apparently none?
  - In order to test the above, we design and implement a series of sequential choice problems in which the subject is asked to make a binary choice in any one and only one of ten subsequent time intervals – one period – while being able to observe all the binary outcomes generated throughout the period. The crust of the problem lies in the fact that those outcomes are generated randomly with a probability of 0.5, and the subject is informed about this fact.
- Does the possibility to observe other subject's decisions influence the decisions?
  - We modify the basic setup above and extend it to a setting, where decisions made by other subjects may be freely observed by the experiment participants.
- Do waiting costs affect the relevant decision processes?
  - In another version of the experiment, we further extend the procedure to include waiting costs, to be incurred from the second period on should the subject choose to postpone their decision.
- Does the inclusion of uncertainty affect the decisions, and if so, in what way?
  - To contrast our findings with the case of there actually being some kind of uncertainty present, we devise and perform for each of the setups mentioned above a series of trials, where the parameter governing the

[*] Ritsumeikan University Department of Economics, Email: izawa@ec.ritsumei.ac.jp
[†] Kinki University Department of Economics, Email: greg@kindai.ac.jp

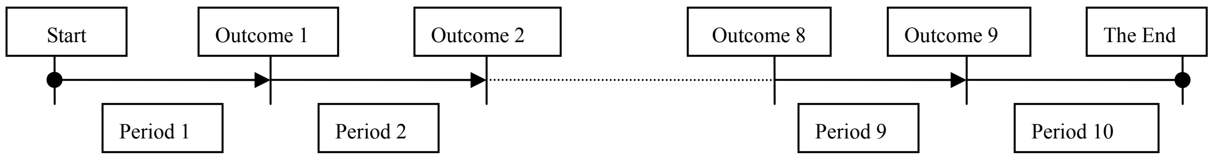
data generating process, i.e. the probability of one of the binary outcomes, is unknown. Hence, by observing subsequent outcomes, the subjects are able to learn and update their estimates so as to improve the quality of their information.

2. The Experimental Design and Procedures

Simply put, a subject's task in this experiment is to predict the outcome of the next movement of a risky stock price. On the computer screen, a series of outcomes is displayed; the outcomes are binary and they are either [U↑], meaning a rise in the stock price, or [D↓], meaning a price drop. For each session, the probability of the price rising is fixed for all ten periods. This probability is either known to the subjects to be equal to 50% or it is unknown to the subjects – on the computer screen it only says "During this session, the probability of the price rising is X". The probabilities are independent across sessions and across different experiment treatments. Furthermore, while the probabilities are exactly the same for all participants, the data is generated independently for each workstation. This is crucial as having access to others' data amounts to being able to observe a larger sample of the data.

The participants are to make their decisions, framed in the language of stock trading, during any one of the ten intervals they choose before a price movement outcome. If the subsequent outcome matches their prediction, the subject wins 10 points; otherwise they lose 10 points. The subject also loses 10 points if they fail to make a decision throughout the session.

A typical session can be adequately illustrated as in the graph below.



The experiments are carefully designed so that all monitors display the relevant information simultaneously. Each period lasts for 15 seconds so that one entire session is 2 minutes and 30 seconds long. Each of the four experiments described below is comprised of 6 sessions, with 10 second breaks between subsequent sessions. After all 6 sessions of an experiment have been terminated, the final score for that experiment is automatically calculated and displayed.

Four treatments of the experiment were conducted as enumerated below. Before each of the treatments, a trial session was conducted to familiarize the subjects with the experimental setup and procedures.

I. Experiment "O" – Isolated Decisions

Subjects are completely isolated from each other. There is no information regarding other participants' actions nor are there any costs associated with delaying the decision-making. The data-generating processes for the 6 sessions are as follows: (The numbers represent the true probability of an "up" outcome; the numbers in parentheses are not communicated to the subjects.)

Session #	1	2	3	4	5	6
Prob. [U↑]	(0.25)	(0.5)	0.5	(0.65)	0.5	0.5

II. Experiment "I" – Observed Decisions

In addition to the generated data displayed on each of the monitors participants were facing, there is also information available regarding other subjects' decisions. This information is updated with each new outcome and is displayed in the form of the numbers of subjects who already made "buy" and "sell" decisions. The data-generating processes are as follows:

Session #	1	2	3	4	5	6

Prob. [U↑]	0.5	0.5	(0.7)	(0.65)	0.5	(0.35)

III. Experiment "C" – Waiting Costs

The setup is identical to the one in Experiment "I", except now costs of delaying a decision are positive. Starting from the second period, 1 point is added to the costs if no decision has been made up until the end of the period. The costs are updated automatically and displayed on the computer display. The data-generating processes are as follows:

Session #	1	2	3	4	5	6
Prob. [U↑]	(0.2)	(0.3)	0.5	(0.3)	0.5	0.5

IV. Experiment "N" – Costly Information

The setup is identical to the one in Experiment "I", except now a subject has to pay to have information about others' actions made available to her. In actuality, none of the participants chose to buy any information. Hence, the resulting setup is equivalent to the basic one of Experiment "O".

Basic Data

The experiment was conducted on the 19th and the 20th of February 2009 at the Ritsumeikan University Experimental Economics Laboratory. Altogether 43 subjects recruited among the University undergraduate student population participated in the experiment. Including instructions preceding the experiment, trial sessions, and payment of rewards afterwards, the experiment lasted about 3 hours both days.

3. Results

The crucial quantity we are interested in is the timing of subjects' decisions. We thus report below in both graphical and tabular manner the normalized distribution of periods decisions were taken in all four of the treatments of the experiment. In the figures below, blue bars represent the relevant percentages for sessions with unknown probabilities while the brown bars represent percentages for sessions when the probabilities were equal to 0.5 and known to the subjects. The numbers on the horizontal axis represent subsequent decision periods.

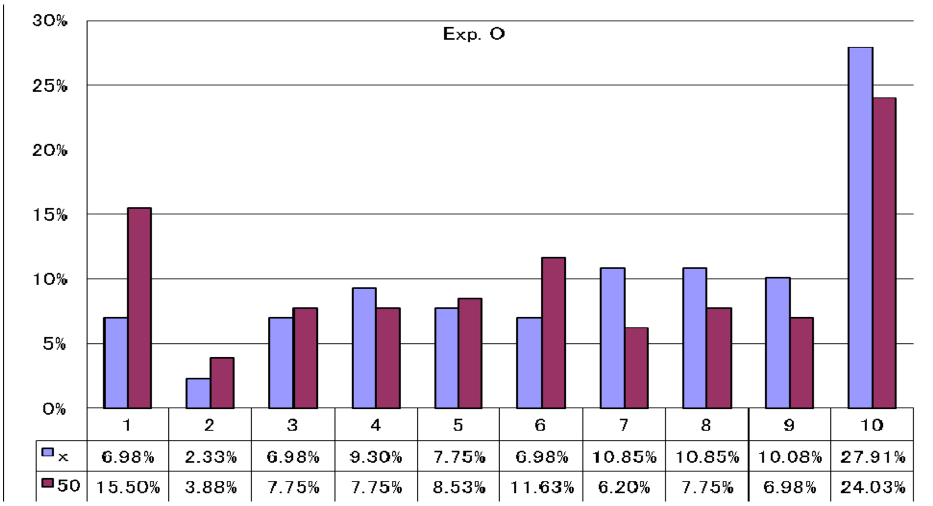


Figure 1: Experiment "O" – Isolated Decisions

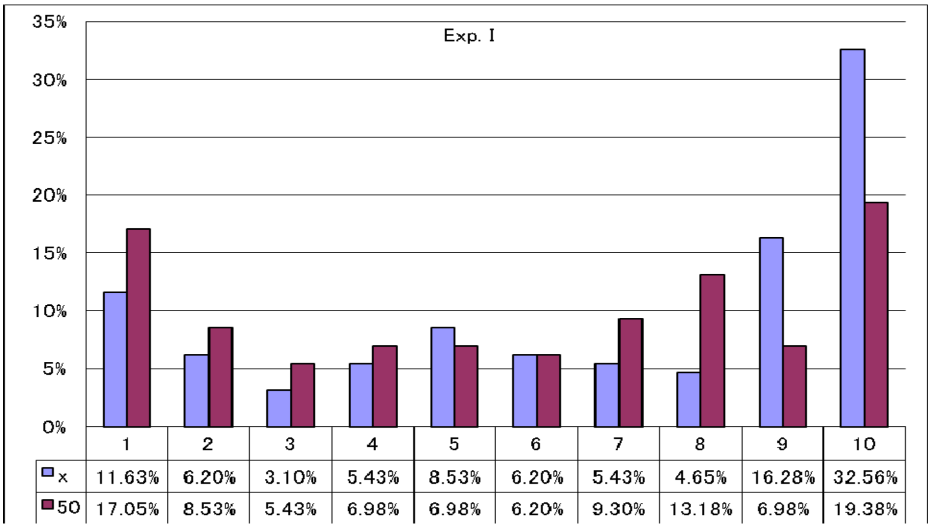


Figure 2: Experiment “I” – Observed Decisions

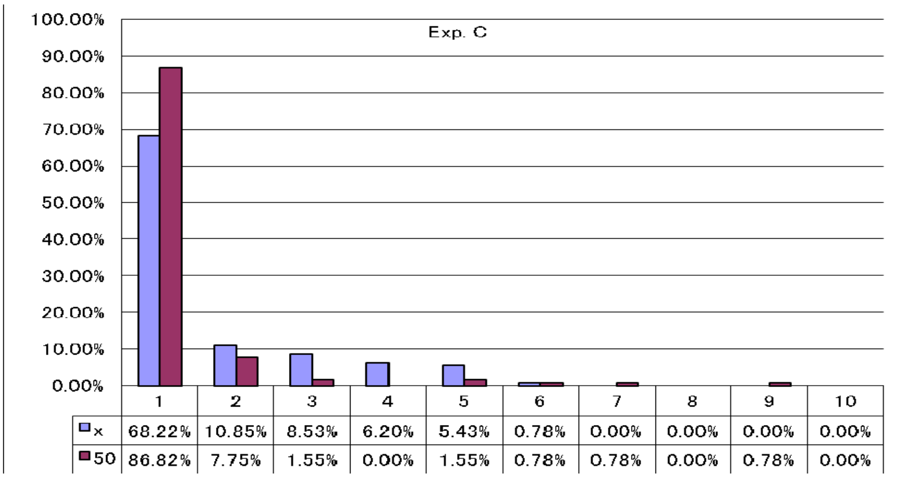


Figure 3: Experiment “C” – Waiting Costs

3.1. Wilcoxon Tests

3.1.1. Averaged data across all treatments – unknown vs. known probabilities:

Avg(x)	5.288372	Wilcoxon V = 563.5	p-value = 0.01442
Avg(50)	4.618217		

3.1.2. Unknown vs. known probabilities in all treatments

O	Avg(x)	6.829457	V = 516	p-value = 0.03405
	Avg(50)	5.984496		
I	Avg(x)	6.821705	V = 533.5	p-value = 0.0447
	Avg(50)	5.775194		
C	Avg(x)	1.720930	V = 148	p-value = 0.004356
	Avg(50)	1.317829		
N	Avg(x)	5.781395	V = 283.5	p-value = 0.4948
	Avg(50)	5.395349		

3.1.3. Unknown probabilities across treatments

O vs. I	Avg(Ox)	6.829457	V = 339.5	p-value = 0.9227
	Avg(Ix)	6.821705		
O vs. N	Avg(Ox)	6.829457	V = 603.5	p-value = 0.02406
	Avg(Nx)	5.781395		
I vs. N	Avg(Ix)	6.821705	V = 546	p-value = 0.002642
	Avg(Nx)	5.781395		

3.1.4. Known probabilities (p=0.50) across experiments

O vs. I	Avg(O50s)	5.984496	V = 389.5	p-value = 0.5731
	Avg(I50s)	5.775194		
O vs. N	Avg(O50s)	5.984496	V = 442.5	p-value = 0.3016
	Avg(N50s)	5.395349		
I vs. N	Avg(I50s)	5.775194	V = 409.5	p-value = 0.3877
	Avg(N50s)	5.395349		

3.2. Randomness Tests

We also tested for the randomness of subjects' responses. When data is generated with 50% probability, i.e. random walk, we would expect the responses to be distributed uniformly – decisions are then made with equal frequency in any of the ten intervals. On the other hand, responses should not be uniformly distributed whenever the data-generating process is unknown.

We use the chi-squared test for given probabilities to test the null hypothesis that probabilities are uniform. Apart from testing for the whole distribution, we also report the results for a truncated frequency matrix where we omit the first and the last interval from the analysis.

Experiment "O"

Interval	1	2	3	4	5	6	7	8	9	10	χ^2	DoF	p-value
Prob=x	9	3	9	12	10	9	14	14	13	36	53.4031	9	2.451e-08
											8.9524	7	0.2561
Prob=0.5	20	5	10	10	11	15	8	10	9	31	39.7597	9	8.4e-06
											5.6923	7	0.5761

When in isolation, responses are not significantly different from random (uniform), both for known and unknown probabilities.

Experiment “I”

Interval	1	2	3	4	5	6	7	8	9	10	χ^2	DoF	p-value
Prob=x	15	8	4	7	11	8	7	6	21	42	90.3023	9	1.416e-15
											21.3333	7	0.003307
Prob=0.5	22	11	7	9	9	8	12	17	9	25	27.5116	9	0.001151
											6.7805	7	0.4521

When given the opportunity to observe others’ decisions, responses are random for Prob=0.5 but significantly different from uniform for unknown probabilities.

4. Discussion

Based on the results presented above, we offer some tentative observations and interpretations with a view to answering the questions stated in the introduction.

Firstly, concentrating on the sessions with known 0.5 probabilities, we note that there is no discernible pattern of decision clustering in any of the three treatments without delay costs, apart from the apparent focal point heuristic. In other words, while the subjects should be indifferent with respect to which time interval to choose to decide in, an average (for all three above-mentioned treatments) of more than 21% of the participants delayed their decisions until the last period. Similarly, an average of more than 17% of the subjects made their decisions during the first period available for doing so. This suggests a simple focal-point phenomenon as a possible explanation.

Secondly, for both cases of known and unknown probabilities, given the possibility of observing other subjects’ decisions has apparently no discernible effect on the distribution of individual decision times: the distribution resulting from Experiment “I” does not appear to be significantly different from either of the distributions from Experiments “O” and “N”, the latter being in effect equivalent to the former (as noted above).

Thirdly, introduction of waiting costs has had an immediate and evident effect on subjects’ decisions. For the 0.5 probability cases, almost 87% of the subjects made their decisions during the first period. For the unknown probability cases, about two out of three decisions, or some 68% of them, were made during the first period, whilst almost 88% of the subjects made their decisions during one of the three earliest periods.

Lastly, and perhaps most strikingly, the subjects in our three experiments with no waiting costs involved (“O”, “I” and “N”) did not in general wait until the final period to make their decisions in cases with unknown probabilities. This is in sharp contrast to the case with delay costs: by not waiting until later periods with their decisions the subjects incur costs due to foregone accuracy of their estimate.

A clear-cut conclusion from the above considerations is that while the subjects would go to great lengths in order to avoid any explicit costs – due to delaying their decisions or to paying for extra information – they apparently do not recognize or choose to ignore implicit costs associated with giving up free additional information that would allow them to refine the accuracy of their decisions.