Random Walk Perception and Timing of Decisions in an Interactive Prediction Task Experiment

Greg Mardyła
Hiroshi Izawa

Abstract

We investigate experimentally individual random walk perception biases and the timing of decisions in a simple interactive prediction task. Our design is fairly general and presents a series of sequential one-shot 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' aggregated choices makes it possible to obtain a more precise estimate of the probability distribution governing the outcomes. This setup mimics thus a stylized stock market model 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, the decision being 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 timing of decisions in random environments. Majority of our subjects exhibit a type of premature timing of actions: in tasks where they should optimally learn as much information as possible and wait until the last period to decide, they make their decisions too quickly, incurring excessive decision costs. This happens even when subjects can observe others' choices at no cost whatsoever.

1. Introduction

Although plenty of research on people's ability to perceive and produce random events as well as on ensuing implications for economic behavior has been conducted during the last four decades starting with the pioneering experiments of Tversky and Kahneman (1971) and Kahneman and Tversky (1972), surprisingly little work has focused on the timing of decisions in problems and environments characterized by such random events. The time a decision maker or a firm chooses to act at in a random environment is of great importance in a wide variety of economic contexts, from stock selection and real investment in general to market entry, to name but a few.

It has become apparent that humans are not well-calibrated judges of randomness, nor are we adequately equipped to produce random sequences, even simple binary ones. Given a stream of totally random data, people often "see" patterns or streaks that make them believe the underlying data-generating process to be non-random and – at least to some extent – predictable. The two phenomena representative of these randomness perception biases are the so-called "hot hand effect" and the "gambler's fallacy". Both phenomena, in their de-facto nature opposite to each other, are based on an apparent belief in "the law of small numbers," according to which people misperceive the frequency with which "runs" or "streaks" of short length occur in randomly generated sequences.

Research on this topic really took off with the work of Gilovich, Vallone, and Tversky (1985), who reported that a basketball player's chance of scoring is believed to be greater following a score than following a miss – hence the naming "hot hand". The question of whether a hot hand in basketball is an illusion or not might actually be very difficult to answer, yet there is no doubt a belief in the existence of patterns in truly random data of the "hot hand" type is a serious and pervasive judgment bias. If people believe the probability of the next outcome in a sequence of completely random binary events to be different from half, their decisions and timing thereof might turn out to be seriously flawed. This observation is our primary motivation for the present research and what exactly is the timing of decisions and actions in such situations is the main question we endeavor to answer.

Specifically, the chief objective of this research is to investigate experimentally in a simple binary interactive prediction task the timing of decisions in random-sequence situations particularly conducive to biases in randomness perception. In particular, we are interested in examining the following issues:
(1) Do decision-makers perceive streaks or trends in data when there cannot possibly be any?
In order to test this issue, 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.
(2) Does the possibility to observe other decision makers' choices 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 in real time by the experiment participants.
(3) Do delay costs affect the relevant decision processes and actions?
In an extended version of the experiment, we further revise the experiment design to include waiting costs, to be incurred from the second period on, should the subject choose to postpone their decision.
(4) Does the inclusion of trends in the data-generating processes affect the timing of actions, and if so, in what way?
We devise and perform for each of the setups mentioned above a series of sessions, in which the parameter governing the data generating process, i.e. the probability of one of the binary outcomes, is unknown to the experiment participants. In this setting, by observing subsequent outcomes and potentially, other subjects' choices, decision makers are able to learn and update their estimates thereby improving the actual precision of their information.

The rest of this paper is organized as follows: in the next section, we outline the design of our experiment. In Section 3, we first informally describe the results and then include some statistical analyses of the data. Finally, in Section 4 we relate our results to existing research and offer some concluding remarks.

2. Experiment Design and Procedures

A subject's task in all sessions and all treatments of the experiment is to predict the direction 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↑] "up", meaning a rise in the stock price, or [D↓] "down", 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 – in this case, the message "During this session, the probability of the price rising is X" is displayed. 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 the opportunity to observe others' actions is tantamount to observing a signal about their data and thus 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, only once during a session. If the subsequent outcome matches their prediction（that is, if the next outcome is “up”/”down” after one buys/sells）, 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.

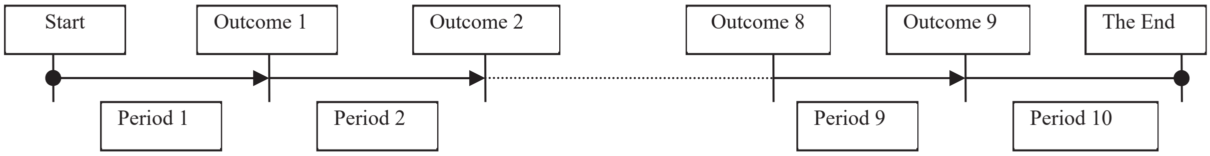


Figure 1: ***Timeline of periods and outcomes within a session.***

The experiments are carefully designed and executed in such a way 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 treatments described below is comprised of 6 sessions, with 10 second breaks between subsequent sessions. Once all the 6 sessions of a treatment have been concluded, the final score for that treatment is automatically calculated and displayed for the subject’s information.

2.1. Experiment Treatments

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

(1) Experiment “O” – Isolated Decisions

Subjects in this treatment are completely isolated from each other. There is no information regarding other participants’ actions nor are there any costs associated with delaying the decision. The data-generating processes for the 6 sessions are as in the table below. The numbers represent the true probability of an “up” outcome; the numbers in parentheses are not communicated to the subjects.

Table 1: ***Data-generating processes in Experiment “O”.***

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

(2) Experiment “I” – Observed Actions

In this treatment, in addition to the generated data displayed on each of the monitors participants were facing, there is also information regarding other subjects’ decisions freely available. This information is updated with each new outcome（after each subsequent period passes）and is displayed in the form of the numbers of subjects who already made their “buy” and “sell” decisions. The data-generating processes are as follows:

Figure 2: *An example screenshot of the actual task for the simplest version of the experiment.*

Table 2: *Probabilities in Experiment "I".*

Session #	1	2	3	4	5	6
Prob. [U↑]	0.5	0.5	(0.7)	(0.65)	0.5	(0.35)

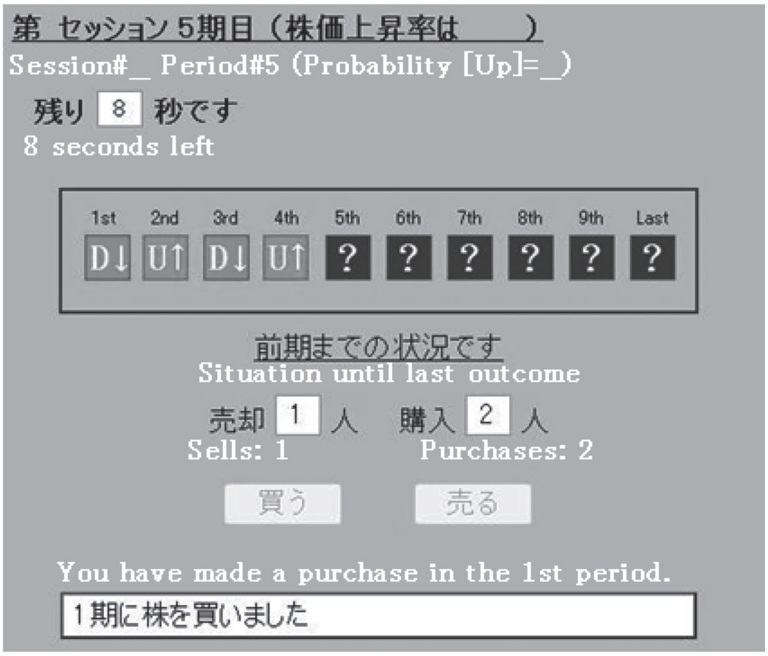


Figure 3: *A screenshot of the actual task for the version where others' decisions are observable at no cost.*

(3) Experiment "C" – Waiting Costs

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

Table 3: ***Probabilities in Experiment "C".***

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

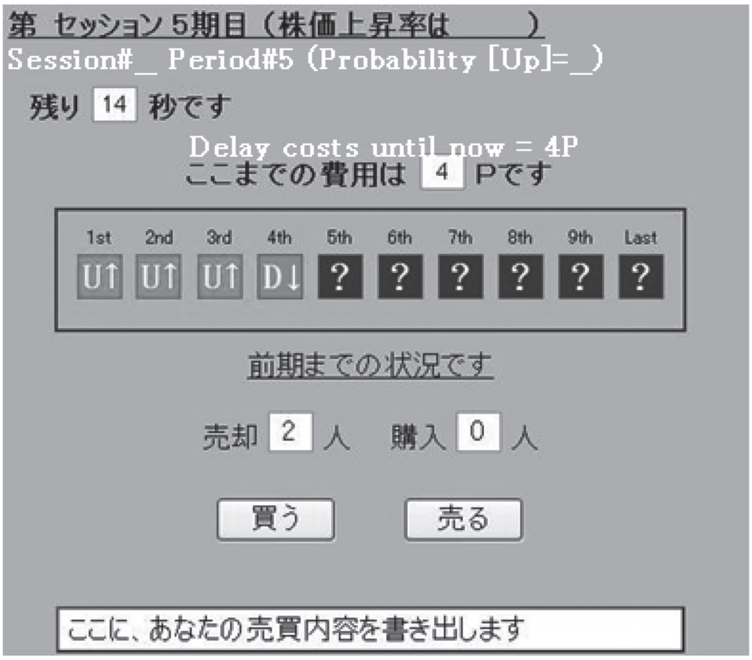


Figure 4: ***A screenshot of the actual task for the experiment version where others' decisions are observable and there are decision delay costs.***

(4) 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. As it turned out, none of the participants chose to buy any information. Hence, the resulting setup is equivalent to the basic one of Experiment "O" and we thus regard it as a "back-to-base" type of setting.

Table 4: ***Probabilities in Experiment "N".***

Session #	1	2	3	4	5	6
Prob.［U↑］	0.5	(0.7)	(0.8)	(0.35)	(0.8)	(0.65)

2.2. 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 to the subjects (in cash directly afterwards), the experiment lasted about 3 hours each day. The average payment, which included a 1000yen show-up fee, was around 3300yen.

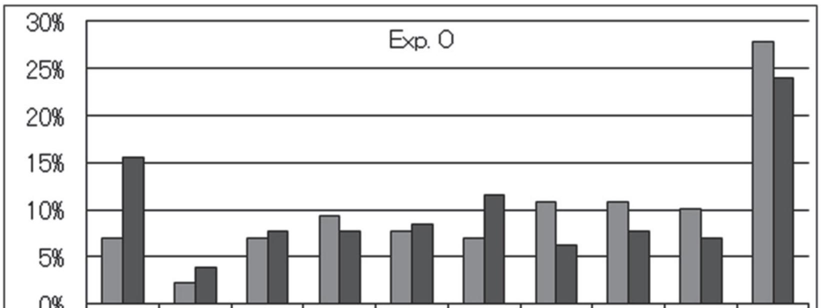
3. Experiment Results

We first present the result of our experiment in an informal graphical manner and then proceed with statistical analyses.

3.1. Graphical Summary of 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 for all four of the treatments of the experiment. In the figures below, blue bars represent the relevant percentages – that is, the percentage of subjects who chose to make their decisions in a particular period – for sessions with unknown probabilities while the purple 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.

3.1.1. Experiment "O" – Isolated Decisions



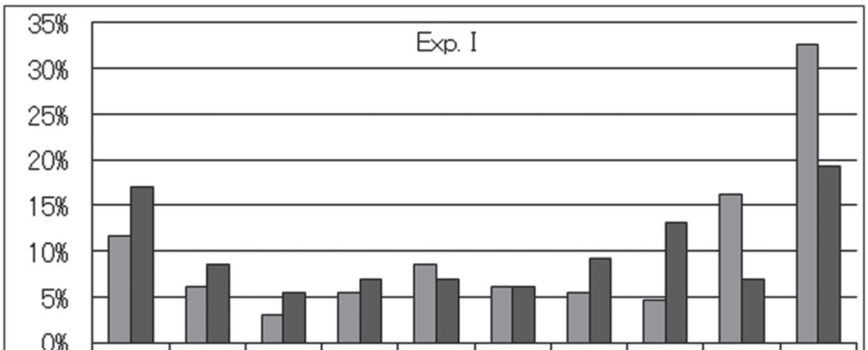
	1	2	3	4	5	6	7	8	9	10
×	6.98%	2.33%	6.98%	9.30%	7.75%	6.98%	10.85	10.85	10.08	27.91
50	15.50	3.88%	7.75%	7.75%	8.53%	11.63	6.20%	7.75%	6.98%	24.03

Figure 5: *Timing of Actions in Experiment "O" – Isolated Decisions*

These results are rather surprising; in particular the distribution of blue bars is not what we would have expected. Since the probability governing the data generating process is unknown, the optimal strategy in this case is to update the equal prior probability as one observes the outcomes appearing on the monitor screen and decide whether to "buy" or "sell" in the last, tenth, period. This is clearly not what is happening here – only some 38% of participants act in the last or the one but last period, while the remaining 62% act prematurely, unnecessarily incurring costs due to foregoing additional acquisition of free information.

As for the three sessions when the probability was 0.5 and known, the purple bars are distributed more or less as we would have expected: there is no discernible pattern of decision clustering. With no explicit costs, it does not matter during which period the decision is made, nor does it matter what the action decided upon actually is.

3.1.2. Experiment "I" – Observed Actions



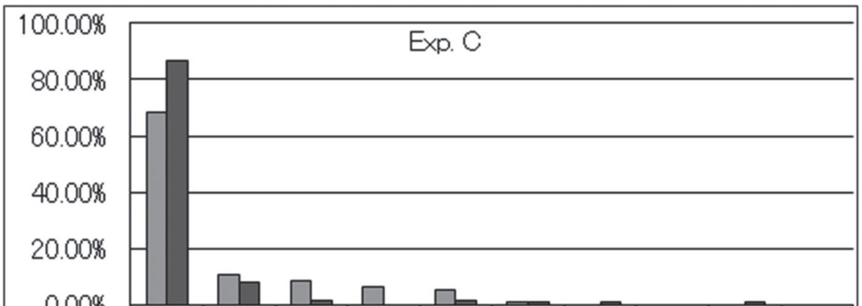
	1	2	3	4	5	6	7	8	9	10
x	11.63	6.20%	3.10%	5.43%	8.53%	6.20%	5.43%	4.65%	16.28	32.56
50	17.05	8.53%	5.43%	6.98%	6.98%	6.20%	9.30%	13.18	6.98%	19.38

Figure 6: *Timing of Actions in Experiment "I" – Observed Actions*

As was the case for the previous treatment, the experiment results are somewhat puzzling here, as well, the more so perhaps given the design of the present treatment. Theoretically, there is no incentive to not wait until the last period to make one's decision here – this feature is common with the previous treatment. What is more, this incentive should ideally be strengthened by the observation of the absence of action on the part of other participants – information regarding others' actions is available in real time and is free to observe. Yet, more than half（some 51%）of the participants chose to act earlier than in the ninth period in the three sessions when the data-generating process is unknown from the outset. While this finding is somewhat closer to what would have been expected from rationally acting decision-makers（49% of subjects act in the last two periods）than in the previously described treatment（where 38% of subjects did）, it still represents a striking departure from optimal decision timing.

Again, our experiment participants seem to have correctly recognized the random nature of data for the three sessions with probability known and equal to 0.5 in this treatment – distribution of the timings of decisions does not appear to exhibit any unexpected regularities in those sessions.

3.1.3. Experiment "C" – Delay Costs

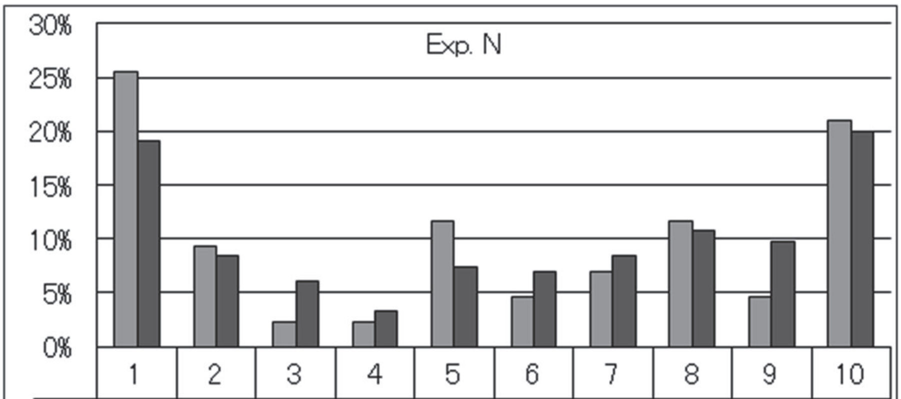


	1	2	3	4	5	6	7	8	9	10
x	68.22	10.85	8.53%	6.20%	5.43%	0.78%	0.00%	0.00%	0.00%	0.00%
50	86.82	7.75%	1.55%	0.00%	1.55%	0.78%	0.78%	0.00%	0.78%	0.00%

Figure 7: *Timing of Actions in Experiment "C" – Delay Costs*

In contrast to the results of previous two treatments, incorporation of explicit delay costs has a clear and well-understood effect on the timing of subjects' decisions. For the probability equal to 0.5 sessions, almost all of the participants (around 94%) chose to act in the first two periods. For the unknown probability sessions also, vast majority (around 79%) of participants acted in the first two periods, with some two out of three participants acting in the first period. Thus, the saliency of delay costs has a direct effect on the decision-makers timing of actions. However, while for the known probability case immediate action is optimal, it is not necessarily so for the unknown probability case. Even taking into account the small sample size of available data, it still is advantageous to delay one's action for a few sessions while updating the prior – majority of the participants exhibit premature timing in this instance, as well.

3.1.4. Experiment "N" –Costly Information



	1	2	3	4	5	6	7	8	9	10
×	25.58	9.30%	2.33%	2.33%	11.63	4.65%	6.98%	11.63	4.65%	20.93
50	19.07	8.37%	6.05%	3.26%	7.44%	6.98%	8.37%	10.70	9.77%	20.00

Figure 8: *Timing of Actions in Experiment "N" – Costly Information*

As mentioned earlier, the setting utilized in this treatment involved a possibility to purchase information about others' decisions. Since no participant actually chose to do so, the resultant decision task is ex-post rendered equivalent to the basic setting in treatment "O". Nonetheless, since the monitor display in treatment "N" differed from that in treatment "O" and since any given subject could not have possibly known whether others had in fact bought any information (including their own), we take particular care when drawing any conclusions with respect to within experiment data based on the results of this treatment. The results themselves are similar to those obtained in the first, basic treatment described above. Participants still do not delay their decisions and actions long enough in the unknown probability treatments. Moreover, in many more instances the subjects act even earlier now, with one out of four actions being taken in the first period. While it points to an apparent absence of learning on the one hand, this pattern of premature "trigger-pulling" should most probably be attributed to the subjects' cognitive resources still being under the influence of the previous "explicit costs" treatment.

3.2. Statistical Tests

We first look at how our experiment participants responded to random data, and then at how random those decisions actually were.

3.2.1. Timing of Decisions in the Aggregate.

We first aggregate the results of all four experiments and apply the Wilcoxon test to check whether the timing of decisions is different depending on whether the data-generating process is known to be a 0.5 probability binary independent process（50）or it is unknown（x）. We would expect the average response time to be later for the unknown probability（x）data. While this is indeed what we find as summarized in the table below, the difference between the two averages is not overwhelming.

Table 5: *Averaged data across all treatments – unknown vs. known probabilities.*

Avg（x）	5.288372	Wilcoxon V = 563.5	p-value = 0.0144
Avg（50）	4.618217		

3.2.2. Unknown vs. Known Probabilities within Treatments.

The next quantities to look at are the decision timing points for the four treatments implemented. The results for the first two treatments – that is, when decisions were made in isolation and when the other subjects' aggregated decisions were observable at no cost – are concordant with the cumulative results reported above: The average response timing is later in treatments "O" and "I" for unknown probabilities; the difference, while significant at the 5% level, is not strikingly large. The timings of responses in treatment "C", in which explicit waiting costs were incorporated, are also in concord with the cumulative results. As for the final treatment "N", surprisingly, there is no significant difference between decision times for the two data-generating processes.

Table 6: *Averaged data within treatments – unknown vs. known probabilities.*

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 = 0.00436
	Avg（50）	1.317829		
N	Avg（x）	5.781395	V = 283.5	p-value = 0.4948
	Avg（50）	5.395349		

We next turn to across-treatment data analysis.

3.2.3. Known Probabilities（p=0.50）across Experiments

Because the data-generating process is totally random and the subjects are aware of it, we would expect no significant differences between responses across treatments. Indeed, this is what we find, as reported in the table below. We interpret these results as evidence that our subjects did understand the task they were facing and were aware of the random nature of data being unraveled on their monitor screens.

Table 7: *Between treatments comparison –known probabilities.*

O vs. I	Avg（O50）	5.984496	V = 389.5	p-value = 0.5731
	Avg（I50）	5.775194		
O vs. N	Avg（O50）	5.984496	V = 442.5	p-value = 0.3016
	Avg（N50）	5.395349		
I vs. N	Avg（I50）	5.775194	V = 409.5	p-value = 0.3877
	Avg（N50）	5.395349		

3.2.4. Unknown Probabilities across Treatments

Contrary to the case of known probabilities, we might expect significant deviations between respective decision timings when the underlying probability distribution is unknown. What we find though, is no apparent difference between subjects' decision-making patterns when deciding in isolation (“O”) and when deciding while observing others' aggregated decisions (“I”). Moreover, as can be verified in the table below, our subjects made their decisions much earlier in the final (“N”) experiment than in either of the first two treatments. This finding suggests the question of whether our subjects' decisions were “too random” to begin with. We turn to this issue in the next section.

Table 8: ***Comparison of action timing between treatments – unknown probabilities.***

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 = 0.00264
	Avg (Nx)	5.781395		

3.3. Testing for Randomness of Actions

We thus also tested for the randomness of subjects' responses in the first two basic treatments of the experiment. 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 theoretically 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. The aim here is to exclude any potential focal-point effects. In fact, as we report in the following subsections, whenever we do not exclude the extreme observations (the 1st and the 10th periods), the results point to a significant difference between the observed frequency and the expected count of decision timing distributions. We consequently concentrate in our reporting and analyses of the ensuing results on the truncated data, i.e. we only take the eight center observations into account.

3.3.1. Experiment “O”: Isolated Decisions

When deciding in isolation, responses are not significantly different from random (uniform), both for known and unknown probabilities. While for the known probabilities case this is what we would have normally expected, the distribution of decisions is apparently too random for the case of unknown probabilities (for the truncated data only, obviously).

Table 9: ***Randomness of responses in the isolated decisions experiment.***

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	53.40	9	2.45e-08
											8.95	7	0.2561
Prob=0.5	20	5	10	10	11	15	8	10	9	31	39.75	9	8.4e-06
											5.69	7	0.5761

3.3.2. Experiment "I": Observed Actions

When given the opportunity to observe others' decisions, responses are random for Prob=0.5 but significantly different from uniform for unknown probabilities, which is precisely what we would expect from optimizing decision-makers. Still, as in other cases, many of our subjects made their decisions much earlier than we would have imagined. We offer some thoughts as to the reasons behind this outcome in the next section.

Table 10: *Randomness of responses in the "observed actions" experiment.*

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	1.41e-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

4. Conclusions

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 no auto-correlation between data when probabilities are known to the subjects to be equal to 0.5, 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（i.e. when there potentially is auto-correlation between data）, 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.

Faced with random data, the waiting time before a decision is made and action taken, while being the main variable we concern ourselves with, is of significant importance in numerous areas of interest, including but not limited to investment and real options theory. As mentioned in the introduction, surprisingly little research has been done on the subject. One recent experimental study that addresses this problem is Oprea, Friedman, and Anderson（2009）. In their irreversible investment experiment, subjects also do not wait long enough before plunging into action, similarly to our results. If given（ample）opportunity to learn, participants in their study do become more patient as they solve more problems, but convergence to optimum is slow and incomplete. "Premature investment", as the authors themselves call the phenomenon, cannot be explained away with（risk/time）preference-based explanations, they claim. Given that the timeframe in their experiment is much shorter in our simple setting, these theories clearly cannot premature actions observed in our experiment, either. What is more, Oprea, Friedman, and Anderson（2009）deliberately induce impatience by design. Since we do not employ any devices to that effect, the more surprising our findings are. The simple "one-shot decision" design we employ does not leave much room for learning opportunities and we thus cannot realistically expect any learning to occur, which it does not.

As it is difficult to provide a convincing explanation for biases in the perception of randomness such as the hot hand effect or gambler's fallacy, we refrain from speculating on the potential causes of suboptimal timing of decisions and actions observed in random environments. A clue to such causes might lie perhaps in the evolution theory paradigm. Recent work by Wilke and Barrett（2009）points to the possibility that the hot hand effect might be the default way we humans perceive random events. Their argument essentially attributes the hot hand effect to humans' evolved response to a world in which it is the norm that items of interest, such as food or mates, come in "clumps". If it is indeed the case then neither randomness nor stochastic correlation or the like are humans' default modes for dealing with data. Perhaps then, waiting with a decision until all potentially valuable information is made available is not the default option, either.

References

Gilovich T, Vallone R, Tversky A（1985）The hot hand in basketball: On the misperception of random sequences. Cognitive Psychology 17:295–314.

Kahneman D, Tversky A（1972）Subjective probability: A judgment of representativeness. Cognitive Psychology 3:430-454.

Oprea R, Friedman D, Anderson ST（2009）Learning to Wait: A Laboratory Experiment. Review of Economic Studies 76: 1103-1124.

Tversky A, Kahneman D（1971）Belief in the law of small numbers. Psychological Bulletin 76:105–110.

Wilke A, Barrett HC（2009）The hot hand phenomenon as a cognitive adaptation to clumped resources. Evolution and Human Behavior 30: 161-169.

（まるでわ　ぐれぐ　Faculty of Economics, Kindai University）
（いざわ　ひろし　Faculty of Gastronomy Management, Ritsumeikan University）