Preliminary Draft. Do not quote or circulate without authors' permission.

Knowing when to stop and make a choice, an experimental investigation.

Ambroise Descamps*
Sébastien Massoni
Lionel Page

Queensland University of Technology

Abstract

We propose an optimal sequential sampling approach to analyse the accumulation of costly information in a risky environment. The nature of the decision marker's problem is twofold: what is the optimal stopping time, and what decision to take. Using sequential analysis tools commonly employed by cognitive scientists, we derive the optimal behaviour of a risk neutral Bayesian agent, depending on the sampling cost, the potential loss or gain incurred, and the sampling horizon. We propose a laboratory experiment to test the theoretical predictions of our model. Specifically, participants can pay a unit cost to receive binary signals about the nature of a risky prospect. Our design conveniently allows the cost of information and the nature of the prospect (gain or loss) to vary. We find that subjects substantially deviate from the optimal strategy in a systematic manner: when information is relatively expensive, participants *oversampled* information, leading to lower payoffs. We find also that participants tend to learn with time and improve their sampling behaviour, without converging to the optimal strategy.

*Corresponding author: ambroise.descamps@qut.edu.au

1 Introduction

The axiom of *completeness*, introduced in early courses on the theory of consumer choice, is often seen as benign and rarely the object of much discussion. Faced with two goods A and B, the consumer is assumed to be able to choose between the two. The axiom does not allow for the consumers not to know if they prefer A or B. Research in psychology and marketing has raised serious questions about this "benign" axiom being an accurate or useful representation of consumer behaviour. In many situations, consumers can be seen as "constructing" their preferences on the fly as choices situations appear in front of them (Lichtenstein and Slovic, 2006). Intuitively, the axiom of completeness conflicts with a lot of situations in our everyday life experience. Who has never been undecided when having to choose whether to buy a given good (car, holiday, house), or to choose between two of such goods (two cars, two different holidays, two houses)?

Our own experience, as well as research in psychology, conflict with a naive application of the axiom of completeness. Recent theoretical research in decision making at the cross-road of cognitive psychology and economics provide a way to solve this problem and extend the economic framework to situation where customers discover over time which option they prefer. It now considers situations where consumers are uncertain about their preferred option because they are not sure which one is most likely to bring them the most satisfaction. Faced with such situations, consumers often have the possibility to acquire additional information on the likely satisfaction they will get from each option. They can stop acquiring information whenever they have enough confidence in the superiority of one option relative to the other. Such situations generate an "optimal sequential sampling" problem: the sampling of information is sequential and the consumer would want to stop sampling when he has just the right amount of information.

In the present study we experimentally investigate how such models can help us understand how people make every day life decisions. We design a situation where individuals have to choose between two options. They are not aware of the payoffs associated with each option. They can choose one of the options at any time or alternatively wait for additional (costly) pieces of information which help them get a better idea of the likely payoffs of each option. The design of the experiment allows us to predict the optimal sequential sampling strategy and to assess whether participants are able to approximate it. Overall, we find that participants deviate significantly from the optimal strategy. When given enough time participants tend in several cases to learn and get closer from the optimal strategy, though we do not observe convergence to the optimal strategy within the time frame of the experiment. We observe a tendency to oversample the information when it is relatively costly: compared to the optimal strategy, participants tend to acquire too many signals. They do make less mistake as a consequence but the gains involved do not compensate for the costs accumulated from sampling information.

As early as 1945 Wald developed a formal framework to find the solution to the question "when should the decision maker stop looking for additional information and make a decision". Recent research in cognitive science suggests that the brain seems to approximate the optimaly strategy generated by Wald's theory when solving simple choice tasks in a fraction of a second (Bogacz, Brown, Moehlis, Holmes, and Cohen, 2006). In somewhat of a twist, such models suggest that preferences over different options may actually be constructed in real time and in an optimal way given the information available and the cost of acquiring more of it. Beyond brain processes and split seconds decision, one may naturally wonder how much such a framework can explain decision making for choices which take place over some tangible durations, from several seconds to several days.

By investigating this question, the present study contributes to two strands of literature. First it contributes to the growing interest in understanding how economic decision are made over time. For many years, decision time was ignored by economist as an irrelevant variable. Recently, economist have turned their attention to the time dimension of the decision making process. In experimental economics, Rubinstein (2007) investigated how decision time (or rather response time) is associated with cognitive reasoning in game situations. In the context of risky choices, Moffatt (2005) found that longer decision time seems to signal greater cognitive cost expanded to depart options which are close in expected utility. Over the recent years, drift diffusion models have been imported from cognitive science to model economic decisions (Webb, 2013; Fehr and Rangel, 2011; Krajbich, Oud, and Fehr, 2014; Caplin and Martin, 2015). They give a structure to the time dimension of the decision process. Noticeably, Ryan Webb, has shown how these models of optimal sampling naturally give rise to random utility type models very similar to those use by microeconomists (Webb, 2013). The randomness of choices was introduced in random utility models for its ability to explain the patterns of observed choices Rieskamp, Busemeyer, and Mellers (2006). The origin of this randomness was assumed to come from unobserved utility or cognitive errors. Models of optimal sampling are able to give a rigorous foundation to this randomness element. They suggest it comes from the intrinsic nature of the information received in the sequential sampling process. By doing so they open the door to the large research on random choice models to consider the time dimension whereby the choice process takes place and how this dimension can influence the nature of the choice randomness. The present study contributes to this strand by looking at sequential sampling situations in economic decisions. In such situations, the cognitive processes used may be substantially different than those engaged in the short choice task from cognitive science studies. The present study investigates how much the sequential sampling framework and its optimal solution can be used to model real economic decisions.

Second, this paper contributes to the literature comparing the predictions of dynamic models of decision making with actual human behaviour. A large part of this literature has risen to test search models which have emerged in labour economics. Job search has been described as an optimal stopping problem where an agent sequentially screens wages offers (Stigler, 1962; McCall, 1970). A large number of studies have investigated human behaviour in search situations where options appears sequentially. Rapoport and Tversky (1966, 1970) proposed the first search experiments. Subjects sequentially receive offers and have to choose whether to accept one of them or receive a new one at a fixed cost. They find that most subjects behave according to the theoretical predictions. Cox and Oaxaca (1989) propose an explicit experimental test of the job search literature, and observe behaviour remarkably fitting that of risk neutral and averse agents. More recently, experimental studies have investigated stopping behaviour in continuous time experiments. Oprea, Friedman, and Anderson (2009) study an optimal stopping problem in a financial setting. They analyze the behaviour of an agent that can incur a fixed cost to seize an irreversible investment opportunity. Their main result is that despite poor performance at the start of the experiment, the subjects learn "intuitive heuristics" to approximate the optimal behaviour. In a follow-up experiment, Viefers and Strack (2014) find that subjects are actually not always consistent and adopt a behaviour suggesting they anticipate regret. This paper extend this type of research by investigating behaviour in an important and though under studied setting: sequential sampling.

The remainder of the paper is as follows: Section 2 presents the conceptual framework of the optimal sequential sampling problem. Section 3 presents the experimental design, Section 4 presents the results and Section 5 concludes.

2 Conceptual framework

2.1 Optimal Sequential Sampling: the General Case

The problem of information accumulation can be described as a sequential choice among several possibilities. At each stage of the problem, the decision maker (DM) has the option to take an action in a finite set or acquire more information. Taking the wrong action and getting information are both costly. As described by Arrow, Blackwell, and Girshick (1949), using the early work from Wald (1945), the DM is in the shoes of a statistician that has to test two hypothesis and can discriminate between them by accumulating information.

Let's assume that a DM wants to test whether a state of the world (θ) is A or

B. He wants to test the two hypothesis: $H_0 : \theta = A$ or $H_1 : \theta = B$[1]. To do so he can gather signals (X_i) at a unit cost c. The signals are discrete random variables, with different distributions under each state of the world, which we will write as $f_\theta(X_i)$. The wealth of the DM depends on the choice he makes at time n, u_n, on the nature of the state of the world.

After the accumulation of n signals, we observe a sequence $X_1^n = (X_1, ..., X_n)$. Let $P(\theta = B) = \pi$, be the probability of the state of the world being B. We choose to present the problem in a case where the action of the DM leads to a potential loss. The case of a gain leads to a symmetrical example and yields the same results. If he makes the wrong decision the DM can incur a potential loss:

$$L(\theta = i; u_n = j) = \begin{cases} 0 & \text{if} \quad i = j = A, B \\ L_0 & \text{if} \quad i = A; j = B \\ L_1 & \text{if} \quad i = B; j = A \end{cases}$$

The total loss from stopping at time n is : $L_n(\theta; u_n; X_1^n) = L(\theta = i; u_n = j) + cn$. Where cn is the total amount spent on signal gathering. Using Bayes' rule and the law of total probability, we can write π_n, the probability of the state of the world being B conditional on the observed sequence:

$$\begin{aligned} \pi_n =& P(\theta = B | X_1^n) \\ =& \frac{P(X_1^n | \theta = B) P(\theta = B)}{P(X_1^n)} \\ =& \frac{P(X_1^n | \theta = B) P(\theta = B)}{P(X_1^n | \theta = B) P(\theta = B) + P(X_1^n | \theta = A) P(\theta = A)} \end{aligned}$$

As signals are i.i.d. and $f_B(X_i) = P(X_i | \theta = B)$, we get[2]: $P(X_1^n | \theta = B) = \prod_{k=1}^{n} f_B(X_k)$. Hence:

$$\pi_n = \frac{\pi \prod_{k=1}^{n} f_B(X_k)}{\pi \prod_{k=1}^{n} f_B(X_k) + (1 - \pi) \prod_{k=1}^{n} f_A(X_k)} \tag{1}$$

[1]The notations are borrowed from Tartakovsky, Nikiforov, and Basseville (2014)

[2]
$$\begin{aligned} P(X_1^n | \theta = B) =& P(X_n \cap X_{n-1} \cap ... \cap X_1 | \theta = B) \\ =& P(X_n | \theta = B) \times P(X_{n-1} | \theta = B) \times ... \times P(X_1 | \theta = B) \end{aligned}$$

Let $G^{st}(\pi_n) = min\{L_1\pi_n; L_0(1-\pi_n)\ \}$ be the expected payoff from stopping the information and $\mathbb{E}[G^N_{n+1}(\pi_{n+1}|\pi_n)]$ the expected payoff from sampling. The sampling decision at every period is made by comparing these two quantities.

Therefore we can write the DM's objective at every period as a maximization problem:

$$\begin{cases} G^N_n = & min\{G^{st}; \mathbb{E}[G^N_{n+1}(\pi_{n+1}|\pi_n)] + c\} & \text{if } n \leq N \\ G^N_N = & G^{st}(\pi_n) & \text{if } n = N \end{cases}$$

Note that in practice, computing the sampling expected payoff is often not possible analytically and requires the us of numerical methods[3]. Figure 1 plots a typical behavior of expected payoffs.

There exist two solutions to the equation $\mathbb{E}[G^N_{n+1}(\pi_{n+1}|\pi_n)] + c = G^{st}(\pi_n)$, that is that make DM indifferent between sampling and stopping.

An is the solution of the equation $\pi_n \leq 1/2$, below which the decision maker chooses the state of the world A, and Bn is the solution of the equation if $\pi_n \geq 1/2$, above which he chooses B. These thresholds determine the sampling region for the decision maker.

A standard way to model accumulation of information is to use a log-likelihood ratio, $Z_k = log(f_B(X_k)/f_A(X_k))$, so that the accumulated information at time n is modeled by a sequential probability ratio test $\lambda_n = \sum_{k=1}^{n} Z_k$. The use of this log-odd ratio allows to summarize information using a simple additive form. Indeed, each piece of information has the same weight in this space, as the products from the conditional probabilities are linearized.

Now that we derived the solutions for the cross section, we can describe the evolution of the process throughout the draws. Note that[4]:

$$\pi_n = \frac{\chi e^{\lambda_n}}{1 + \chi e^{\lambda_n}} \Leftrightarrow \lambda_n = log\left(\frac{\pi_n}{(1-\pi_n)\chi})\right) \tag{2}$$

So that we can rewrite the decision the log-likelihood space:

$$\begin{cases} \text{Gather more information:} & \text{if} \quad a_1(n,N) \geq \lambda_n \geq a_0(n,N) \\ \text{Choose State of the World A:} & \text{if} \quad \lambda_n \leq a_0(n,N) \\ \text{Choose State of the World B:} & \text{if} \quad \lambda_n \geq a_1(n,N) \end{cases}$$

[3]For instance, in the case of a binary signal, for a given X_1^n, there are 2 possible sequences X_1^{n+1}, one of which overlaps with another X_1^n. That means that there are n possible sequences for X_1^{n+1}. This leads to $\sum_{k=1}^{n} n$ possible sequences at after n signals. If $n = 100$, that the expected payoff has to be computed for 5050 cases.

[4]A proof is provided in appendix

Where:

$$\begin{cases} a_1(n,N) & = log\left(\dfrac{B_n}{(1-B_n)\chi}\right) \\ a_0(n,N) & = log\left(\dfrac{A_n}{(1-A_n)\chi}\right) \end{cases}$$

A typical plot of these boundaries is plot on Figure 2. It is worth noting that

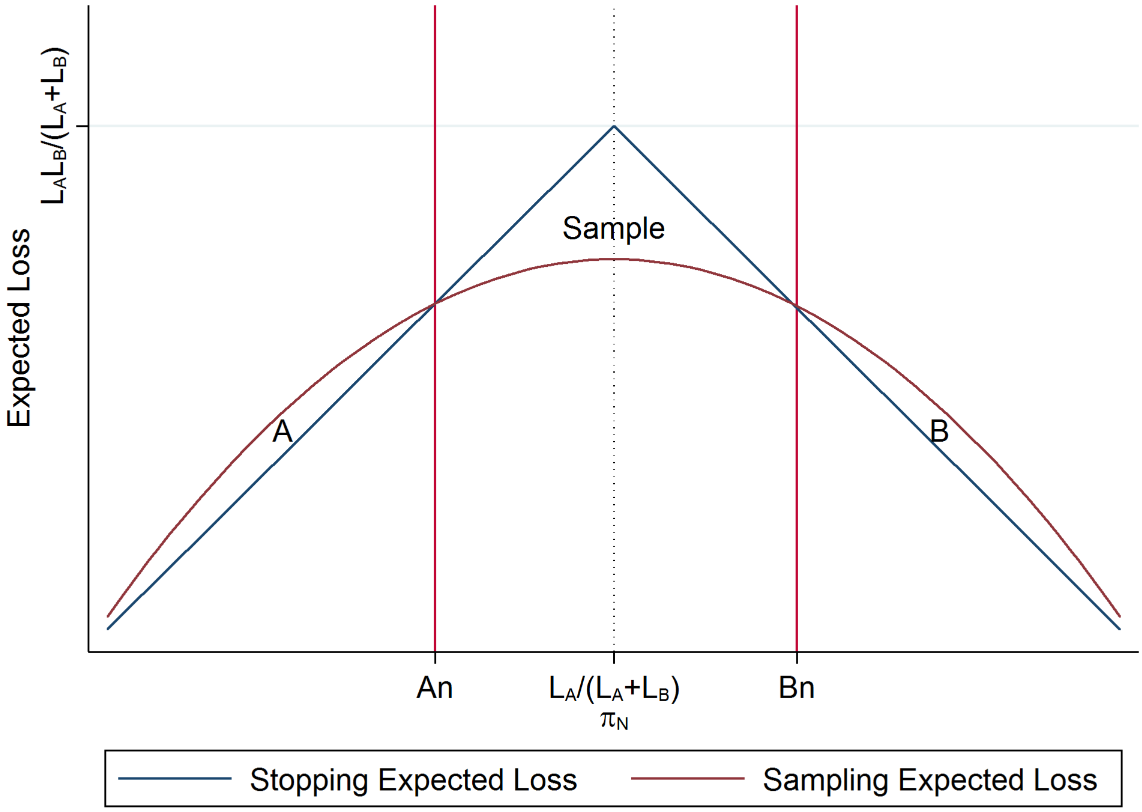

Figure 1: Typical plot of expected losses. The blue lines represent the stopping expected loss depending on the level of prior (π_n). The expected loss peaks when uncertainty is the highest, and is the lowest when the DM is the most certain about of the state of the world. The parabola represent the expected payoff from sampling. Both are equal in two points, A_n and B_b, which are the solutions to $\mathbb{E}[G^N_{n+1}(\pi_{n+1}|\pi_n)] + c = G^{st}(\pi_n)$.

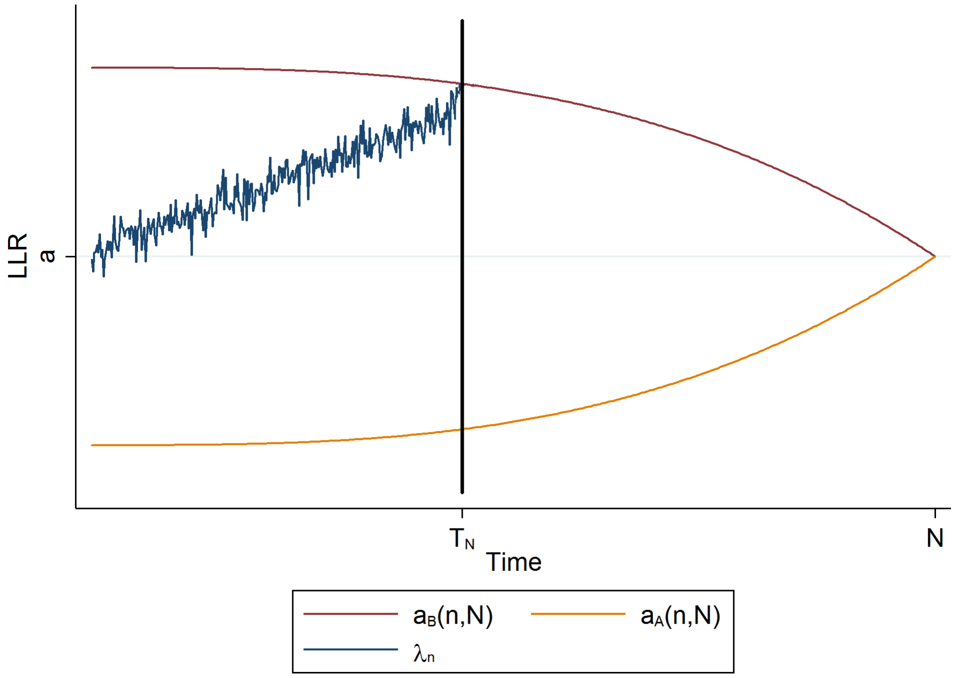

Figure 2: This plots the typical plot of the decision boundaries in the log odds space. The upper boundary is the limit after which the DM will choose state of the world B, and the lower boundary makes the DM choose A. The blue line depicts the typical behaviour of a log-likelihood ratio.

2.2 Optimal Sequential Sampling: The Special Case of Binary Signals

To study behaviour in a sequential sampling, we design an experiment with two states of nature (urns with different proportion of white and black balls) and where decision makers receive binary signals (draw of a ball). We describe here the optimal behaviour in such a special case of the precedent model. For illustration, we use the parameters used in the experimental design.

Let's assume that both states of the world are as likely ($\pi = 0.5$), and that the signal is binary ($X_i \in \{0, 1\}$), so that X_i follows a Bernouili distribution. Under state of the world A, $X_i \sim \text{Bern}\,(0.4)$, and under state of the world B, $X_i \sim \text{Bern}\,(0.6)$. Thus $f_B(X_k) = 0.6^{X_i} \times 0.4^{1-X_i}$.

Let $Y = \sum_{k=1}^{n} X_i$, be the number of times a "1" signal is received. We can write

$\prod_{k=1}^{n} f_B(X_k) = (0.6^Y \times 0.4^{n-Y})$ and $\prod_{k=1}^{n} f_A(X_k) = (0.4^Y \times 0.6^{n-Y})$. Which yields:

$$\pi_n = \frac{(0.6^Y \times 0.4^{n-Y})}{(0.6^Y \times 0.4^{n-Y}) + (0.4^Y \times 0.6^{n-Y})}$$

In order to compute the expected payoffs we need to derive all the possible combinations of signals. Let's assume that $n = 100$. For ease of representation, we assign a value of +1 to a if $X_i = 1$ and -1 if $X_i = 0$. This is represented in Figure 3.

For each dot on Figure 3, we are able to assign the probability of the sequence of draws using the formula we derived for π_n. This is displayed in Figure 4.

In our experiment we chose $L_0 = L_1 = \$20$, and various cost of sampling ($c = \$0.1, \$0.5, \$1$). Using the priors we computed we are able to derive the expected loss from stopping ($G^{st}(\pi_n)$).

Noting that at time $n = N$ the only possible choice is to stop, we can compute the expected loss from sampling at $n = N - 1$: $\mathbb{E}[G^N_{N-1}(\pi_N|\pi_{N-1})] = G^{st}(\pi_N)$. Using backward induction from $n = N-1$, we can derive the expected payoff from sampling at any period.

Hence, we are able to find the boundaries, A_n and B_n, that make the DM indifferent between sampling or not by solving the equation $G^{st}(\pi_n) = \mathbb{E}[G^{n+1}_n(\pi_{n+1}|\pi_n)] + c$. We plot these boundaries in the signal and prior spaces (Figures 3 and 4).

We are able to translate the equivalent of all these variables in the log-likelihood ratio space. This is depicted in Figure 5. This allows us to conveniently represent the accumulation of information by a DM. The additive form of information in a log-odd space can be seen in Figure 5 as the distance between each dot appears constant.

Figure 6 shows the evolution of a typical information accumulation in the log likelihood ratio space. In this example, the optimal behaviour is to stop after 5 draws as λ_n crosses the frontier.

Set of all Possible Signals Combinations

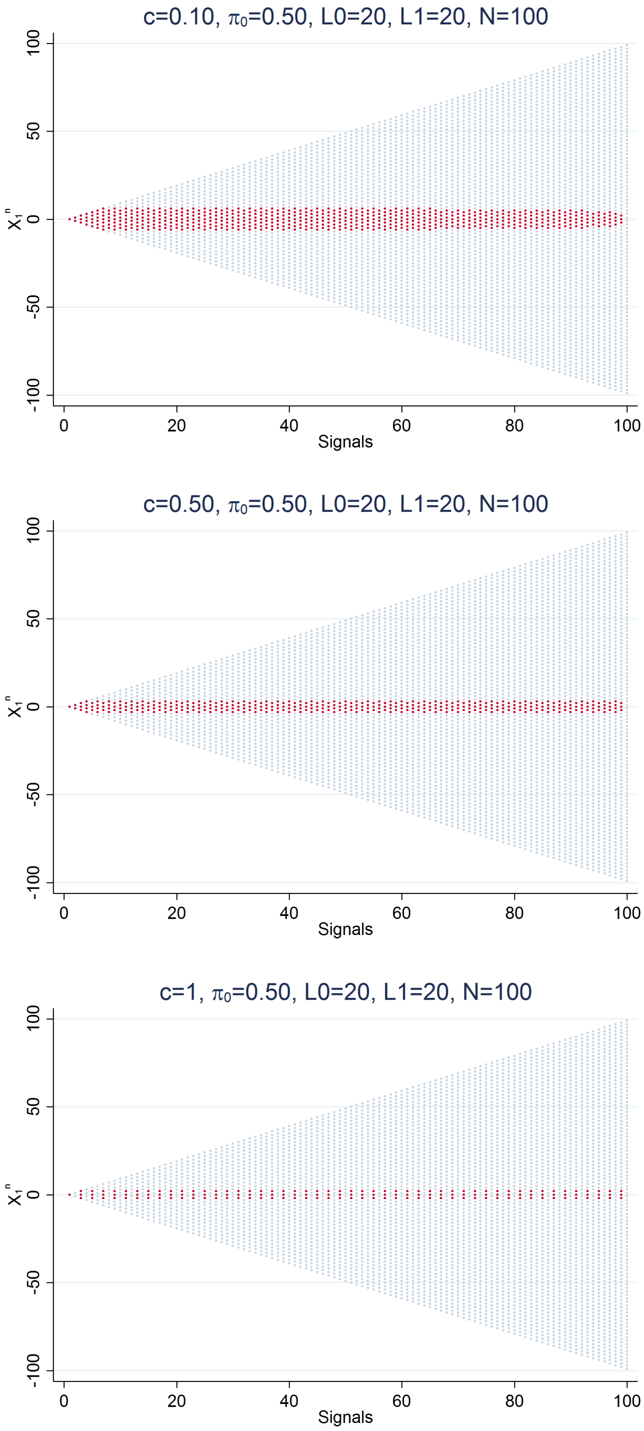

Figure 3: The blue triangle depicts all the possible combination of binary signals that can be accumulated when $n = 100$. The red part represents the combination of signals for which the DM should continue to sample information, given the parameters corresponding to each panel.

Probability of each State of the World for Signal Sequence.

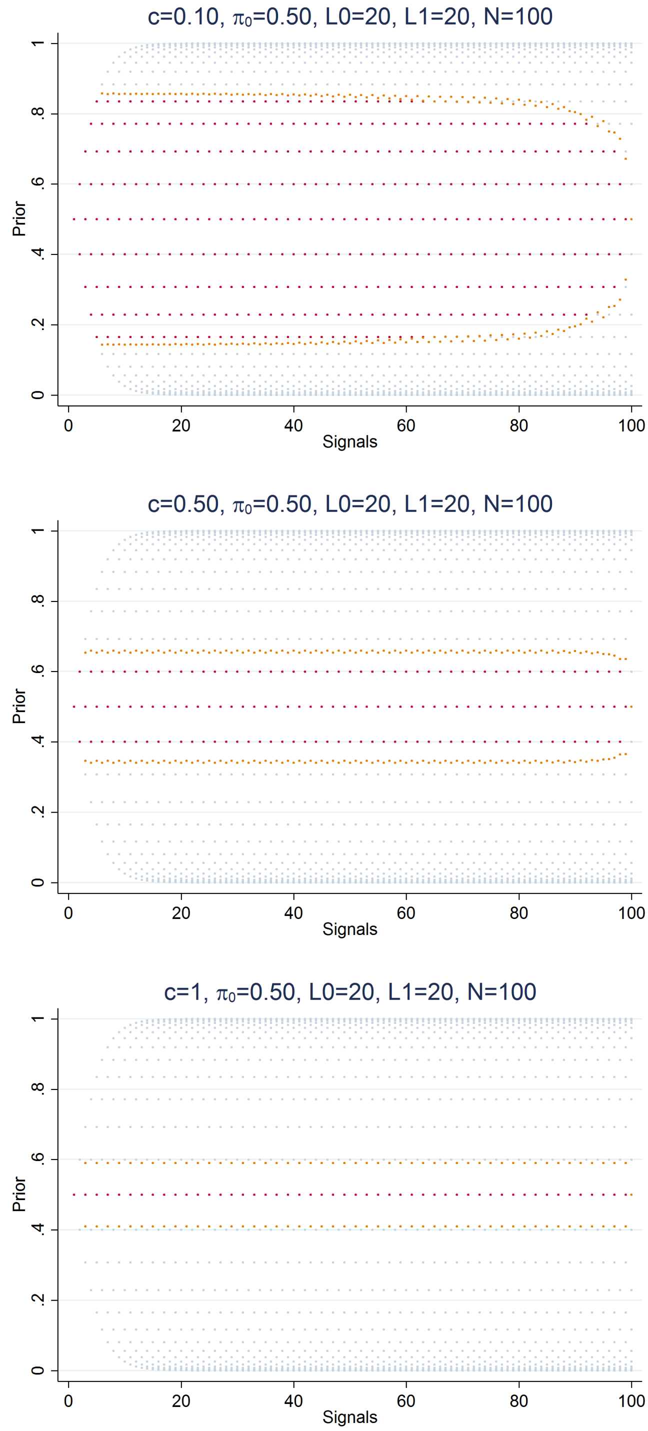

Figure 4: This depicts the probability of the state of the world being B, given all the accumulated signals (π_n). The orange dots represent the values that make the DM indifferent between sampling or not ($A_n \& B_n$). The red dots represent the level of beliefs that are too uncertain to reach a decision.

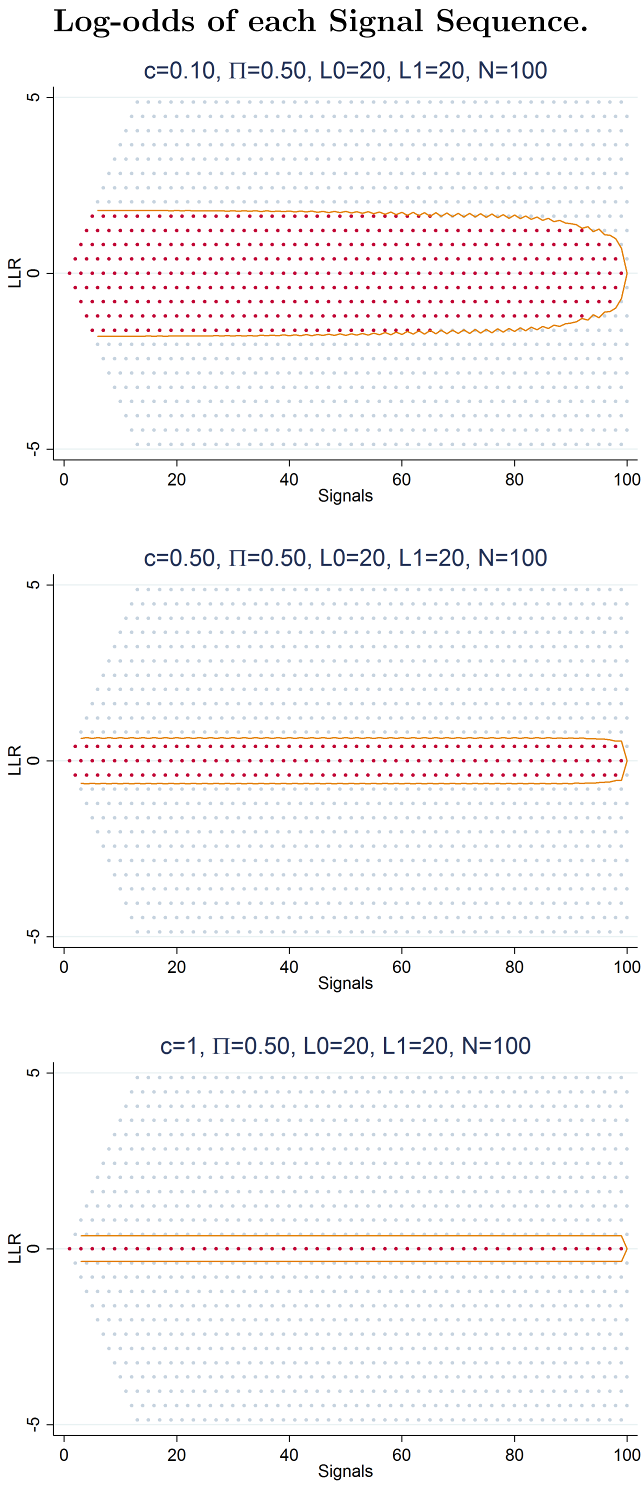

Figure 5: This depicts the log-odds, that is the relative likelihood of each state of the world for a given sequence depicted in Figure 3. a_n and b_n represent the log-odds that make the decision make indifferent between sampling or not. The read area represent the log odds that are too uncertain to reach a decision.

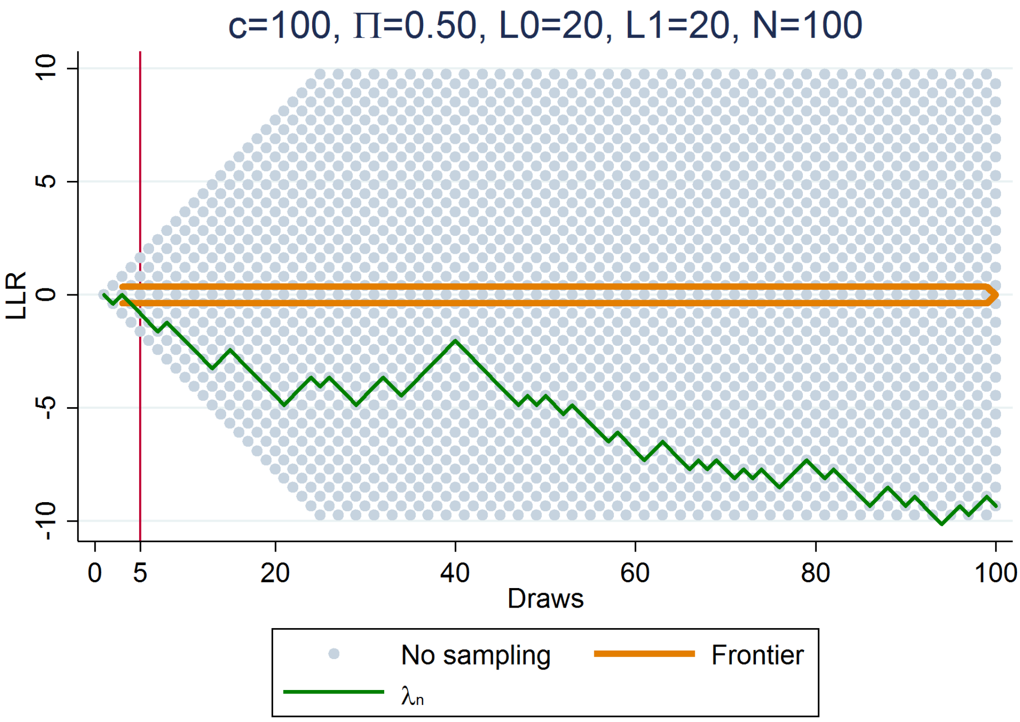

Figure 6: This plots the typical behaviour of a DM following the optimal stopping rule we derived. Given the signals that have been accumulated, the optimal strategy is to stop after 5 signals.

3 Experimental Design

3.1 Testable Hypotheses

This section summarizes the features and predictions of the model that can be tested using a laboratory experiment.

Hypothesis 1. ***Optimality.*** *A risk neutral agent should follow the principles of optimal sequential sampling.*

When facing the parameters from section 2.2, a risk neutral agent should stop acquiring information when the log likelihood ratio of the sequence crosses the frontier that we derived above. As such, we are able to provide a normative analysis of the behaviour observed in our experiment.

Hypothesis 2. ***Comparative Statics.*** *Individuals react in the right direction when the cost of information changes.*

The prediction of our model regarding this change in incentives is unambiguous and intuitive. Because of the high cost of information, the trade-off between making better decisions and getting higher expected loss more stringent. Hence, the higher the sampling cost, the lower the optimal number of draws

Hypothesis 3. ***Neutrality to Framing.*** *Framing the decisions should have no effect on the search process.*

The assumption we made to get our predictions is that the agent is risk neutral. Hence, presenting the search problem as a way to avoid a loss or to make a gain should have no impact on the stopping time.

Hypothesis 4. ***Learning.*** *If the participants do not behave optimally, giving them some feedback on their performance should allow them to learn how to approximate the equilibrium.*

If **Hypothesis 1** is violated, we can expect the participant to improve throughout the experiment. Studies have shown that participants are able to learn to approximate optimal behaviour in mathematically demanding problems, such as equilibrium bidding in a double auction market (Friedman and Rust, 1993) or approximating the optimal timing of investment (Oprea, Friedman, and Anderson, 2009). An interesting feature of our model shown in section 2.2, is that it provides intuitive heuristics that can be learned after some practice: for a low cost of sampling, individuals should accumulate at least 5 signals and wait to have 4/5 more in a direction, for a medium cost, they should get at least 3 signals and wait to have 2/3 more in one direction and when the cost is high, they should only acquire 2 signals.

Hypothesis 5. ***Irrelevant Information.*** *When a DM has to start a new search, and a new state of the world is determined by nature, the current decision should not be correlated with past decisions.*

A standard finding about series of decision, is the so called Gambler's fallacy (Tversky and Kahneman, 1971; Rabin and Vayanos, 2010). In sequential decision making, individuals tend to overestimate the alternation rate between alternatives. For instance, in a coin flipping experiment, the sequence "HTHT" seems more likely than the "HHHT" one, even though they are equally probable. Hence, after observing a state of the world A, in the next choice he will have to make and given all the information accumulated, the DM should not choose B more often.

3.2 Design and Treatments

We propose a search experiment where the participant has to decide in what state of the world he is, which will determine his payoff. At the start of the experiment, one of two urns is randomly selected by a computer with a probability $\pi = 0.5$. One of these two urns contains 4 black balls and 6 white balls and the other one contains 6 black balls and 4 white balls. Both are depicted on Figure 7.

The participant has to guess what urn has been selected. He can pay a unit cost to see a ball which goes back, into the urn. The balls are drawn automatically (every 2s) and the participant has to stop the process. If he makes a correct guess, the participant receives $20. The subjects had to take 80 decisions in every session. For each of the decision a new urn is selected, but the share of black and white balls remains the same.

At most we allowed them to draw 100 balls from the urn. To make sure that the participants did not get a negative payoff at the end of the experiment, we give them a $10 endowment.

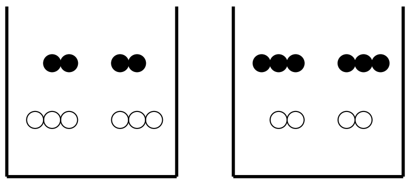

Figure 7: Sketch of Urn A and B.

We propose a design with three levels of cost and two frames to test our hypotheses. By using three different costs (low -$0.10-, medium -$0.50- and high -$1-), we can test **Hypothesis 2** and analyze how participants react to standard economic incentives.

We frame the decision in two ways to test **Hypothesis 3**. In the loss frame we inform the participants that they were all given $20 just by sitting in the room, and that correctly guessing the nature of the urn will allow them to keep this money. On the other hand, in the gain frame we tell them that if they make the right guess, they will get $20. In theory, this should not affect their behaviour.

Using an incentivized risk elicitation procedure (Holt and Laury, 2002) to determine the risk profiles of the participants, we will be able to check **Hypothesis 1**.

By repeating the task multiple times we will be able test **Hypothesis 4&5**. A new urn will be randomly picked by the computer before each round. Studying the evolution of distance to the optimal behaviour will tell us whether agents learn with experience. Finally, as each round will be independent, there should be no correlation between the decisions made by the agents.

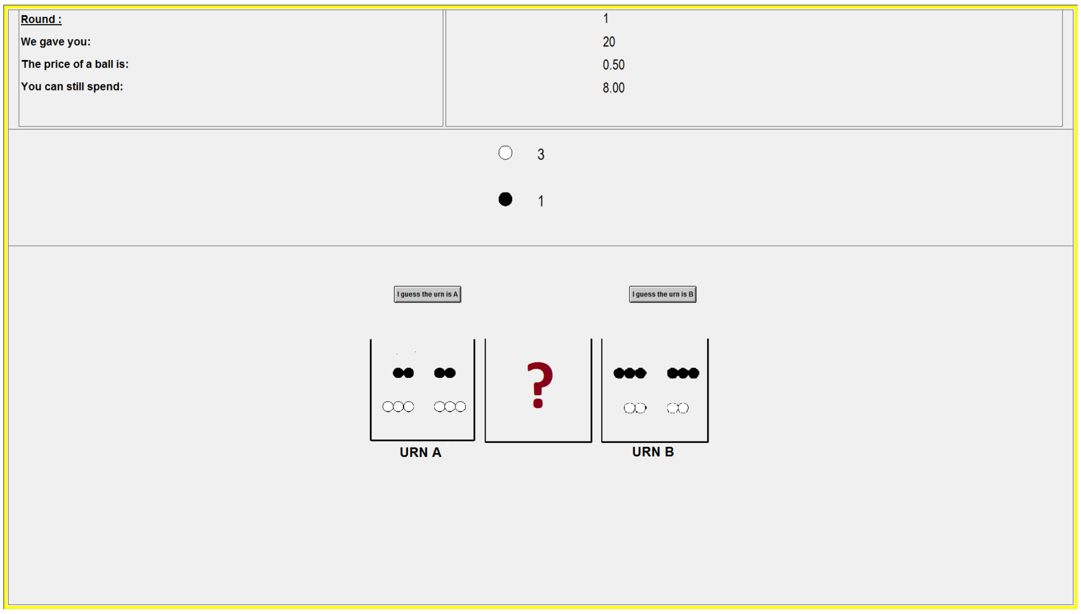

Figure 8: Sample Screen of the search experiment

3.3 Experimental Procedure.

The experimental sessions were conducted at the Queensland Behavioural Economics lab. The participants were students from various faculties at the Queensland University of Technology, and were recruited through ORSEE (Greiner, 2015). In total, 183 students took part in the experiment (mean age 23.5, 60% males). Each session was approximately two hours long.

After the 80 rounds were completed, the participants entered an incentivized risk preferences elicitation task (Holt and Laury, 2002). At the end, the subjects were asked a few demographic questions and had to complete a cognitive reflection test.

The average payoff is $28.6 per round (including a $5 show up fee). One of the round was randomly selected by the computer, and the subjects were given the corresponding payoff.

4 Results.

First we want to investigate whether agents behave optimally in our experiment. Figure 9 displays the share of optimal sampling and deviations from it. Overall, the figure is quite low (8.19%). We derived our model predictions using a risk neutral agent. If we focus on individuals have such risk preferences[5], we observe they have an optimality rate of 7% ($N = 23$), whereas risk averse and risk loving participants respectivly have a rate of 8.39% and 8.45% ($N = 125$ and $N = 34$). An analysis of variance reveals there is no significant difference in behaviour between these three groups ($p = 0.89$). It is worth noting that despite the poor performance with regards to optimality, the success rate in finding the right urn is much higher than it would be if chosen randomly[6]. A visually striking feature of our results appears in Figure 9. The share of undersampling is high when the cost of information is low (76%) whereas oversampling is high for the two other treatments (67% for $c = \$0.5$ and 96% for $c = \$1$).

A similar pattern appears for risk neutral agents: undersampling is high in the low cost treatment and high in the other two treatments, while the rate of optimal behaviour is low. This is depicted in Figure 10.

Result 1. *We find that individuals deviate from optimality, contrary to **Hypothesis 1**.*
***Result** 1.1 The rate of optimal behaviour is low among risk neutral individuals.*
***Result** 1.2 We observe oversampling when information is expensive, and undersampling when it is cheap.*

Second, we want to analyze the response to incentives changes. Figure 11 represents the CDF of stopping times for each level of sampling cost. Visually, the CDFs are ordered according to the sampling cost: the leftmost curve represents the stopping time for the high cost treatment, the middle curve is the distribution in the medium cost treatment, and the rightmost one is the low cost treatment. It appears that each distribution first order stochastically dominates the one corresponding to a cheaper cost. This suggests that participants have reacted to the change in the cost of information as expected. Table 1 summarizes the stopping behaviour for each treatment. We do observe that, when the cost of sampling increases, the average number of balls drawn from the urn decreases. This figure ranges from approximately 5 balls in the high cost treatment, to around 7 balls in the medium cost treatment, and up to 10 in the low cost.

Result 2. *We find support for **Hypothesis 2**. When the cost of sampling increases, the average stopping time decreases.*

[5]Defined as such by our risk elicitation procedure.

[6]72% in the low cost treatment, 65% in the medium cost treatment and 61% in the high cost one

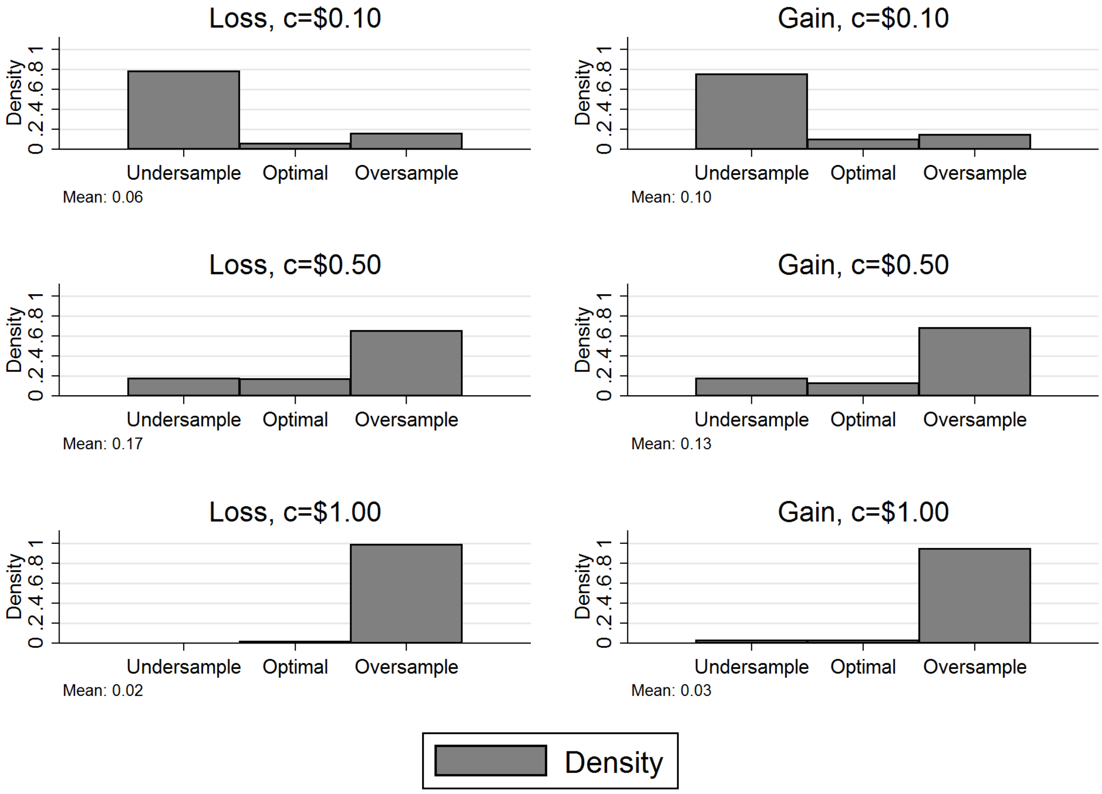

Figure 9: This plots the share optimal behaviour in the whole sample. Undersampling is defined as taking less balls that what is expected at the optimal stopping time, conversly oversampling means that too much information is acquired.

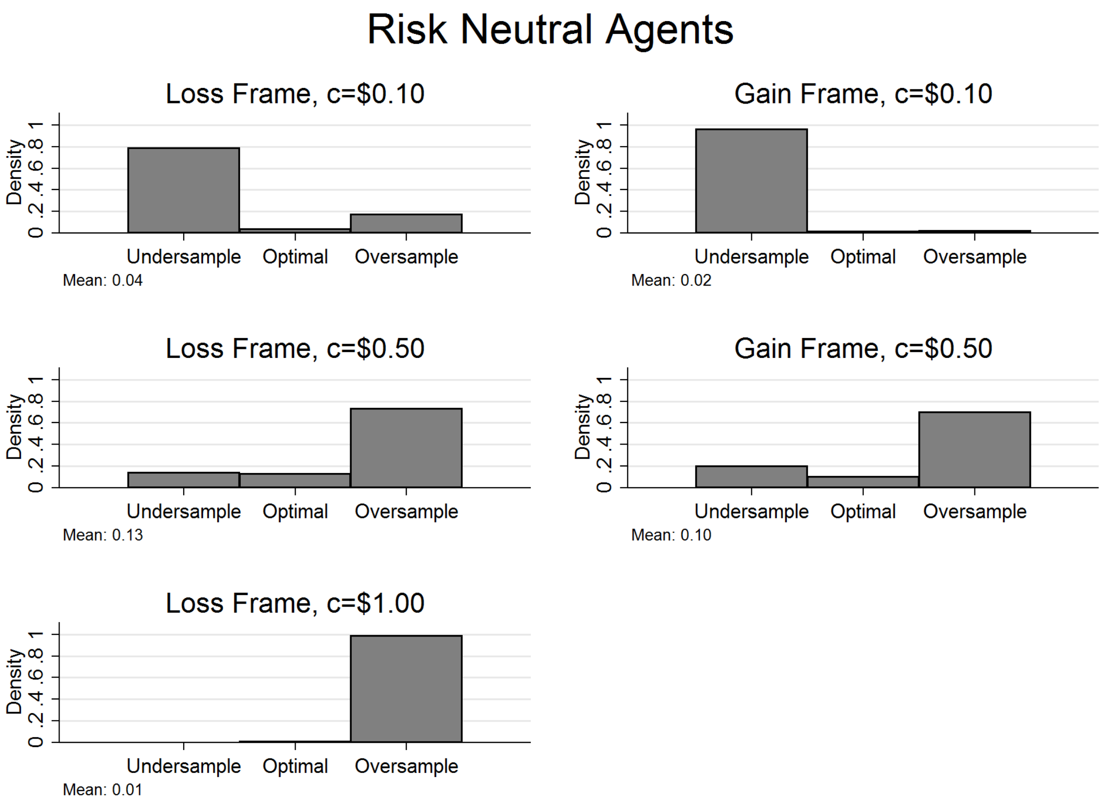

Figure 10: This plots the share optimal behaviour only for the risk neutral agents, according to the Holt and Laury procedure. No participant was categorized as risk neutral in the gain treatment with a high cost.

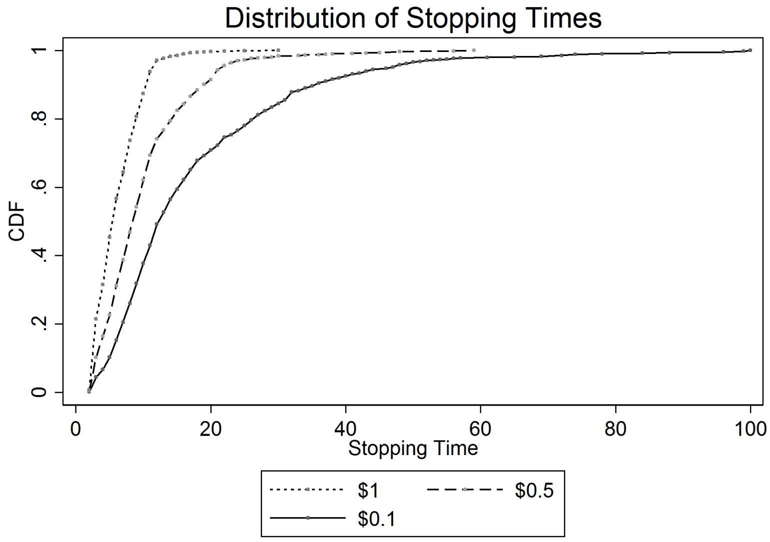

Figure 11: This plots the CDF for each level of sampling cost in our experiment.

Looking at the framing effects, no consistent pattern appears regarding its impact on the stopping times. It appears that it is weakly significantly smaller in the gain frame when the cost of sampling is low ($p = 0.096$). In the medium cost treatment the stopping time is significantly higher in the gain treatment ($p < 0.001$). In the high cost treatment the difference is not different from zero ($p = 0.66$). Overall behavioural biases do not seem to play a role in search problems.

Result 3. *In accordance with* ***Hypothesis*** *3, the framing does not impact the sampling behaviour in our experiment.*

Do individuals learn from their past mistake? One of the main finding in Oprea, Friedman, and Anderson (2009) is that after a learning phase, participants are able to approximate the optimal behaviour in the stopping problem. In our case, we are able to locate how far the subject is from the optimal boundary in the log-odd space. Figure 12 depicts the evolution of the distance at the stopping time, throughout the 80 rounds of the experiment. It appears that, on average, the subjects tend to improve their performance and get closer to the optimal behaviour for the medium and high cost treatments. The oversampling and learning results in the medium and high cost treatments might suggest that the the higher cost makes the information search problem more salient to the participants.

Table 1: Summary Stats of Stopping Time

		Gain	Loss	Theory
10 cts	N	2800	2160	
	Mean	9.88	10.29	10.24
	se	7.35	9.90	8.15
50 cts	N	2560	2240	
	Mean	7.50	6.82	4.37
	se	4.96	4.68	2.22
100 cts	N	2720	2559	
	Mean	5.22	5.19	2.06
	se	2.70	2.63	0.40

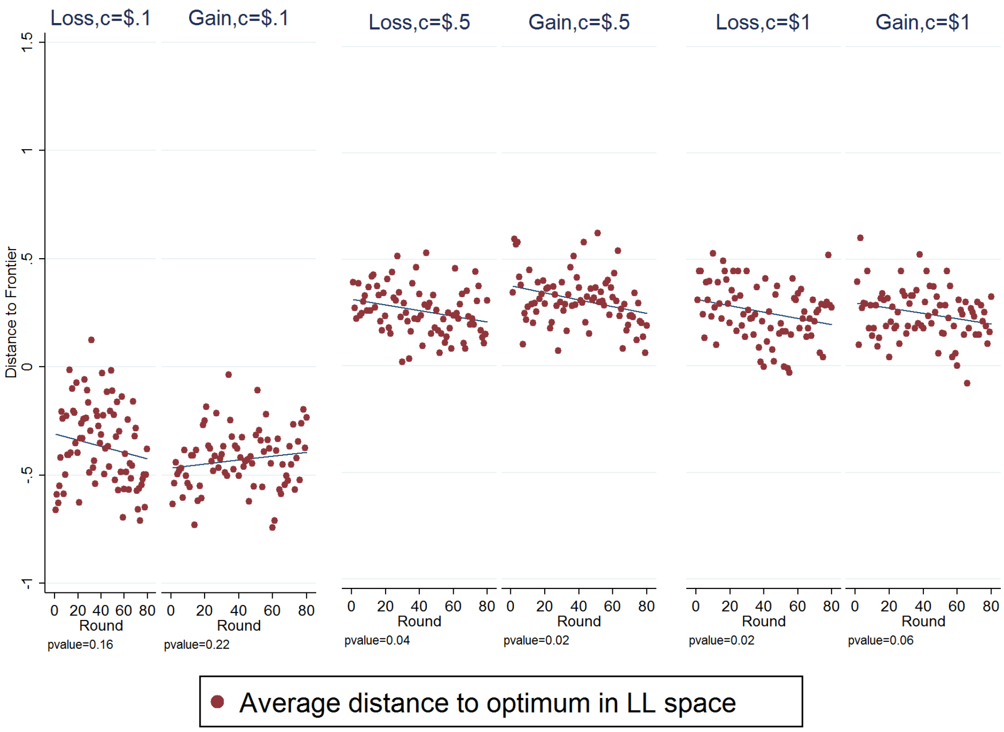

Figure 12: Distance to the optimal frontier.

Result 4. ***Hypothesis 4*** *is partially verified. After some practice, participants in the medium and high cost treatments tend to learn and get closer from the optimal strategy.*

Finally, we use a specification similar to Chen, Moskowitz, and Shue (2014) to test for the Gambler's fallacy:

$$Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + Controls + \epsilon_{it}$$

Where Y_{it} is the decision that has been taken by individual i in round t about the state of the world. β_1 measures the autocorrelation between the decisions that have been take by the agent, that is what share of the current choice is explained by the past decision. As the urns are randomly selected before each round, we should expect $\beta_1 = 0$. $\beta_1 < 0$ would suggest that participants start a round with a prior that they *should* alternate their decisions. We can control for the quality of past decisions, as we, and they, know whether they were right or wrong. We can also control for the intensity of the signal they had when making their mind (X_n^1).

Table 2 and Table 3 plot the result of an OLS estimation of this specification, splitting the sample by level of cost and framing. We observe the gambler's fallacy in each of our six specifications. The coefficient is negative and highly significant in all cases. We observe that the intensity of the effect ($|\beta_1|$) is greater when the cost of sampling increases (from 2% to 7%, $p < 0.05$ in all cases), which is consistent as a higher cost leaves more room for randomness. The behaviour in the gain and loss frames is strikingly similar as we observe a 4% negative autocorrelation ($p < 0.001$). Note that this effect is strictly speaking between a Gambler's fallacy as the autocorrelation is between the decisions that have been taken, regardless of the actual state of the world. Interestingly, the actual state of the world in the past round influences the current decision only when the cost of sampling is the highest and in the loss frame. This effect is of similar magnitude in both treatments ($\beta = -0.02$, $p < 0.05$).

Result 5. ***Hypothesis 5*** *is violated. We find robust evidence of negative autocorrelation in the decision making of our participants.*

Table 2: Testing for the Gambler's fallacy (1)

	0.1\$	0.5\$	1\$
Lag Choice	-0.02*	-0.03**	-0.07***
	(0.010)	(0.010)	(0.011)
Lag Right	-0.01	-0.01	-0.02*
	(0.011)	(0.011)	(0.011)
Value of signals	0.11***	0.13***	0.13***
	(0.001)	(0.002)	(0.002)
Observations	4503	4740	5213
R^2	0.581	0.505	0.383

Standard error in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Testing for the Gambler's fallacy (2)

	Gain	Loss
Lag Choice	-0.04***	-0.04***
	(0.009)	(0.008)
Lag Right	-0.00	-0.02*
	(0.009)	(0.009)
Value of signals	0.12***	0.12***
	(0.002)	(0.001)
Observations	6872	7584
R^2	0.468	0.491

Standard error in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Conclusion

This study has investigated to what extent economic choices between options/goods with uncertain values can be appropriately represented by an optimal sequential sampling process. To do so we have designed an experiment where participants have an imperfect information about the payoffs associated with two possible choices A and B. These choices can be thought as representing goods to purchase or more generally economic option (eg. investments). Participant can choose to wait and collect more information or to stop (at any moment) and select one of the options. As we know exactly the participant's information and the options' payoffs, we are able to determine the optimal behaviour in this situation: the optimal sequential sampling strategy. We are then able to measure whether and how participants deviate from optimality.

Our results are twofold. First, we find that, unlike studies in cognitive studies, participants deviate substantially from the optimal strategy. When sampling is relatively expensive, participants *oversampled*, leading to lower expected payoffs. Second, we find that participants tend to learn with time to improve their sampling strategy. However, in spite a running for a high number of periods (N=80) we do not observe convergence to the optimal strategy. The oversampling of information decreases but persists till the end of the experiment.

Given the widespread nature of such economic decision, this result opens new questions for investigation: First, is this pattern characteristic of typical sequential sampling situations? It is worth noting that such an oversampling of information has also been observed in a study of traders decision to invest in information before trading on an experimental market Page and Siemroth (2015). Second, if such a pattern does characterize a wide range of situations, it will be relevant to investigate whether a behavioural bias explains it and whether people can learn the optimal strategy when they get enough experience (eg. experts).

Appendix

Proof of Equation 2

$$
\begin{aligned}
\pi_n = &\quad \frac{\pi \prod_{k=1}^{n} f_B(X_k)}{\pi \prod_{k=1}^{n} f_B(X_k) + (1-\pi) \prod_{k=1}^{n} f_A(X_k)} \\
\pi_n = &\quad \frac{1}{1 + \frac{(1-\pi)}{\pi} \prod_{k=1}^{n} \frac{f_A(X_k)}{f_B(X_k)}} \\
\Leftrightarrow \quad \frac{1}{\pi_n} = &\quad 1 + \frac{(1-\pi)}{\pi} \prod_{k=1}^{n} \frac{f_A(X_k)}{f_B(X_k)} \\
\Leftrightarrow \quad \frac{1}{\pi_n} - 1 = &\quad \frac{1-\pi}{\pi} \prod_{k=1}^{n} \frac{f_A(X_k)}{f_B(X_k)}
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad \frac{\pi_n - 1}{\pi_n} = &\quad \frac{1-\pi}{\pi} \prod_{k=1}^{n} \frac{f_A(X_k)}{f_B(X_k)} \\
\Leftrightarrow \quad \frac{\pi_n - 1}{\pi_n} \times \frac{\pi}{1-\pi} = &\quad \prod_{k=1}^{n} \frac{f_A(X_k)}{f_B(X_k)} \\
\Leftrightarrow \quad \frac{\pi_n}{\pi_n - 1} \times \frac{1-\pi}{\pi} = &\quad \prod_{k=1}^{n} \frac{f_B(X_k)}{f_A(X_k)} \\
\Leftrightarrow \quad log\left(\frac{\pi_n}{\pi_n - 1} \times \frac{1-\pi}{\pi}\right) = &\quad log\left(\prod_{k=1}^{n} \frac{f_B(X_k)}{f_A(X_k)}\right) \\
\Leftrightarrow \quad log\left(\frac{\pi_n}{\pi_n - 1} \times \frac{1-\pi}{\pi}\right) = &\quad \underbrace{\sum_{k=1}^{n} log\left(\frac{f_B(X_k)}{f_A(X_k)}\right)}_{\lambda_n}
\end{aligned}
$$

References

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949): "Bayes and minimax solutions of sequential decision problems," *Econometrica, Journal of the Econometric Society*, pp. 213–244.

Bogacz, R., E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen (2006): "The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks.," *Psychological review*, 113(4), 700.

Caplin, A., and D. Martin (2015): "The dual-process drift diffusion model: evidence from response times," *Economic Inquiry*.

Chen, D. L., T. J. Moskowitz, and K. Shue (2014): "Decision-making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires," .

Cox, J. C., and R. L. Oaxaca (1989): "Laboratory experiments with a finite-horizon job-search model," *Journal of Risk and Uncertainty*, 2(3), 301–329.

Fehr, E., and A. Rangel (2011): "Neuroeconomic foundations of economic choice-recent advances," *The Journal of Economic Perspectives*, 25(4), 3–30.

Friedman, D., and J. Rust (1993): *The double auction market: institutions, theories, and evidence*, vol. 14. Westview Press.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 114–125.

Holt, C. A., and S. K. Laury (2002): "Risk aversion and incentive effects," *American economic review*, 92(5), 1644–1655.

Krajbich, I., B. Oud, and E. Fehr (2014): "Benefits of neuroeconomic modeling: new policy interventions and predictors of preference," *The American Economic Review*, 104(5), 501–506.

Lichtenstein, S., and P. Slovic (2006): *The construction of preference*. Cambridge University Press.

McCall, J. J. (1970): "Economics of information and job search," *The Quarterly Journal of Economics*, pp. 113–126.

Moffatt, P. G. (2005): "Stochastic choice and the allocation of cognitive effort," *Experimental Economics*, 8(4), 369–388.

Oprea, R., D. Friedman, and S. T. Anderson (2009): "Learning to wait: A laboratory investigation," *The Review of Economic Studies*, 76(3), 1103–1124.

Page, L., and C. Siemroth (2015): "An experimental analysis of information acquisition in prediction markets," *Games and Economic Behavior*.

Rabin, M., and D. Vayanos (2010): "The gambler's and hot-hand fallacies: Theory and applications," *The Review of Economic Studies*, 77(2), 730–778.

Rapoport, A., and A. Tversky (1966): "Cost and accessibility of offers as determinants of optional stopping," *Psychonomic Science*, 4(1), 145–146.

——— (1970): "Choice behavior in an optional stopping task," *Organizational Behavior and Human Performance*, 5(2), 105–120.

Rieskamp, J., J. R. Busemeyer, and B. A. Mellers (2006): "Extending the bounds of rationality: evidence and theories of preferential choice," *Journal of Economic Literature*, pp. 631–661.

Rubinstein, A. (2007): "Instinctive and cognitive reasoning: A study of response times," *The Economic Journal*, 117(523), 1243–1259.

Stigler, G. J. (1962): "Information in the labor market," in *Investment in Human Beings*, pp. 94–105. The Journal of Political Economy Vol. LXX, No. 5, Part 2 (University of Chicago Press).

Tartakovsky, A., I. Nikiforov, and M. Basseville (2014): *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press.

Tversky, A., and D. Kahneman (1971): "Belief in the law of small numbers.," *Psychological bulletin*, 76(2), 105.

Viefers, P., and P. Strack (2014): "Too Proud to Stop: Regret in Dynamic Decisions," *Available at SSRN 2465840.*

Wald, A. (1945): "Sequential method of sampling for deciding between two courses of action," *Journal of the American Statistical Association*, 40(231), 277–306.

Webb, R. (2013): "Dynamic constraints on the distribution of stochastic choice: Drift Diffusion implies Random Utility," *unpublished, New York University.*