

# Learning to hesitate

Ambroise Descamps<sup>1,2</sup>, Sébastien Massoni<sup>1</sup>, and Lionel Page <sup>\*3</sup>

<sup>1</sup>School of Economics and Finance, Queensland University of  
Technology

<sup>2</sup>Oxera Consulting LLP

<sup>2</sup>University of Technology Sydney, Economics Discipline Group

## Abstract

We investigate how people make choices when they are unsure about the value of the options they face and have to decide whether to choose now or wait and acquire more information first. We design a laboratory experiment to study whether human behaviour is able to approximate the optimal solution to this problem. We find that participants deviate from it in a systematic manner: they acquire too much information (when costly) or not enough (when cheap). These deviations costs participants between 10% and 25% of their potential payoffs. With time, participants tend to learn to approximate the optimal strategy.

*Keywords:* search; decision under uncertainty; information; optimal stopping; real option.

*JEL Classification:* C91, D81, D83

---

\*Descamps: ambroise.descamps@qut.edu.au; Massoni: sebastien.massoni@gmail.com; Page: lionel.page@uts.edu.au.

For detailed and constructive comments we gratefully thank Martijn van den Assem, Aurélien Baillon, Sebastian Ebert, Drew Fudenberg, Changxia Ke, Martin Kocher, Morten Lau, Philipp Strack, Susan Thorp, Peter Wakker, Joel van der Weele and Leonard Wolk. We also benefited from comments by seminar and conference participants at the Economic Science Association World Meetings (Sydney), Erasmus School of Economics in Rotterdam, Monash University, University of Technology of Sydney, University of Sydney, University of Melbourne, VU Amsterdam, Griffith University, Australian Meeting of the Econometric Society (Sydney), Australian Conference of Economists (Adelaide), Econometric Society Summer School (Kyoto), Asian Meeting of the Econometric Society (Kyoto). Harriet Smith provided excellent research assistance. The views expressed in this article are those of the authors alone.

# 1 Introduction

In the economic theory of decision under uncertainty (risk or ambiguity) the decision maker has to choose between alternatives while facing an irreducible uncertainty about their final payoffs. However, in a wide range of situations the decision maker has the possibility to reduce the uncertainty she faces by looking for information before making a choice. Typical examples include the following situations: a consumer can search for additional reviews before making a purchasing decision, an agent can look for more information before choosing between different investment opportunities, a doctor can request new tests before making a diagnostic, an inspector can decide to investigate a case further before making a decision. In such situations, the decision maker is confronted with an “optimal sequential sampling” problem whereby she needs to decide when to stop acquiring information about the different alternatives and make a choice between them. As early as 1945, Wald developed a formal framework to find the solution to this question. For a Bayesian decision maker (henceforth DM), the optimal solution is to gather information until the marginal expected value of information is smaller than the cost of sampling information.

Does this model offer a good representation of actual individual behaviour? The empirical evidence is scarce, which is somewhat surprising given the general nature of this framework and the widespread situations where it can be used. In the present study we experimentally investigate whether this theoretical framework helps us understand how people make decisions. We design a controlled situation where individuals have to choose between two alternatives with uncertain payoffs. Before making a choice, they have the opportunity to wait and collect additional (costly) pieces of information which help them get a better idea of the likely alternatives’ payoffs. The design of the experiment allows us to precisely identify the optimal sequential sampling strategy and to assess whether participants are able to approximate it.

Our results are clear, we find that participants deviate in systematic ways from the optimal strategy. They tend to hesitate too long and oversample

information when it is relatively costly and on the contrary to make too quick decisions and undersample information when it is relatively cheap. As a consequence of this deviation, participants lose between 10% and 25% of their potential payoff. However, participants learn to get closer to the optimal strategy over time, as long as information is relatively costly. When the cost of sampling information is high, participants end up approximating quite closely the optimal sampling strategy.

The present study contributes to two strands of literature. First, it contributes to the empirical study of the ability of optimal sequential sampling models to explain human behaviour. The optimal sampling framework has recently attracted a lot of interest to explain human decision making. In psychology, it is now seen as giving a foundation to a wide range of cognitive models (e.g. drift diffusion models) where it has been found to appropriately describe the information accumulation process in the brain. While such models have mainly focused on perceptual choices (Gold and Shadlen, 2002; Bogacz et al., 2006), recent studies have suggested to extend these models to economic choices (Webb, 2013; Fehr and Rangel, 2011; Krajbich et al., 2014; Caplin and Martin, 2015). Optimal sequential sampling models can explain several empirical patterns. For instance, they imply that choices should be partly random due to the random nature of new information. As a consequence they offer a foundation to random utility models (Webb, 2013; Woodford, 2014). Such models can also explain apparently puzzling patterns such as the negative correlation between decision times and quality of decisions (Fudenberg et al., 2015). While the interest for the optimal sequential sampling framework is growing in economics we still know little about its ability to explain actual behaviour. Our study adds to this emerging literature. To the best of our knowledge, it is the first experiment to directly test whether individuals approximate the optimal sequential sampling solution when they hesitate before making a choice.

Second, this paper contributes more broadly to the study of the empirical relevance of real option theory for individual decisions. When having to decide between making a choice now and waiting to get more information, the DM has to compare the cost of collecting the additional information to

the option value of the future information. Such a situation is a type of real option problem. In the real option literature, it has been investigated in the specific case of a decision between making an irreversible investment now or waiting for more information, at the risk of seeing the investment opportunity disappear (McDonald and Siegel, 1986). An empirical test of this theory was proposed by Oprea et al. (2009). They analyse the behaviour of agents who 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 follow-up experiments, Della Seta et al. (2014) find that real option models’ predictions with respect to risk aversion are supported, while Viefers and Strack (2014) find that subjects are not always consistent and adopt a behaviour suggesting regret. Our paper extends this type of research by investigating behaviour in a general case where the DM faces two possible prospects with unknown values and has to weigh the value of collecting additional information against the cost of this information.

The rest of the paper is as follows: Section 2 presents the conceptual framework of the optimal sequential sampling model, Section 3 describes our experimental design, Section 4 presents our results and how participants’ choices deviate from the optimal solution to the problem, Section 5 investigates possible reasons for these deviations and Section 6 concludes.

## **2 Conceptual framework: optimal sequential sampling**

### **2.1 Statement of the model**

When hesitating to make a choice, the DM has the option to take an action in a finite set or to acquire more information.<sup>1</sup> Wald (1945) gives the example

---

<sup>1</sup>This is clearly the case when the DM can access new external information when waiting to make a choice. The framework we describe can also be used in situations where no new external information is available. In such situations, taking more time to engage in cognitive introspection can help reduce the uncertainty about the values of the different

of a statistician who has to discriminate between two hypotheses, and has the opportunity to sequentially acquire more observations. Gathering more data comes at a cost, but as the sample size grows the statistician has a finer knowledge of the underlying data generating process. The optimal sequential sampling (OSS) model formalises this problem to determine the optimal stopping rule.

We model the behaviour of a Bayesian DM that sequentially investigates the nature of a state of the world,  $\theta = \{A; B\}$ , by accumulating some evidence on it. Time is discrete, finite and indexed by  $t \in \{0, \dots, T\} = \mathbb{T}$ . The problem of the DM is twofold, she has to find what state of the world she is in, and when to make this decision. We denote by  $d_t = \{a; b\}$  the set of actions at any given period  $t$ , where choosing  $a$  ( $b$ ) is the right action to take when the true state of the world is A ( $B$ ).

To make up her mind, the DM can gather independently and identically distributed (i.i.d.) binary signals ( $X_t \in \{-1; 1\}$ ). These signals are informative about the nature of the state of world (*i.e.*  $\mathbb{P}(X_t = 1|\theta = B) > \mathbb{P}(X_t = 1|\theta = A)$ ), but come at a unit cost  $c$ . Given all the gathered signals ( $X_1^t = (X_1, \dots, X_t)$ ), we can construct the Bayesian DM's posterior at time  $t$ ,  $\pi_t$ , about the likelihood of state of the world B :

$$\pi_t = \frac{\mathbb{P}(\theta = B) \prod_{k=1}^t \mathbb{P}(X_k|\theta = B)}{\mathbb{P}(\theta = B) \prod_{k=1}^t \mathbb{P}(X_k|\theta = B) + \mathbb{P}(\theta = A) \prod_{k=1}^t \mathbb{P}(X_k|\theta = A)}$$

When the expected utility from stopping is greater than the expected utility from sampling, the DM makes a guess about the nature of the state of the world she is facing. We consider a ‘‘certain difference’’ optimal sequential sampling problem (Fudenberg et al., 2015), whereby the DM knows for sure the difference between the utility of making the right ( $U_1 \in \mathbb{R}$ ) and the wrong guess ( $U_0 \in \mathbb{R}$ ). Let  $U(d_t = i; \pi_t)$  be the expected utility from choosing action  $i$ , given the DM's Bayesian beliefs.<sup>2</sup> When stopping at time  $t = \tau$ , the DM choices by matching previous memories with perceived features of the available choices (Ratcliff et al., 2016).

<sup>2</sup>That is  $U(d_t = a; \pi_t) = \pi_t U_0 + (1 - \pi_t) U_1$  if  $d_t = a$  and  $U(d_t = b; \pi_t) = \pi_t U_1 + (1 - \pi_t) U_0$

chooses the action that maximises her expected utility:

$$G(\tau, \pi_\tau) = \max_{i=a,b} U(d_\tau = i; \pi_\tau) \quad (1)$$

The sequential nature of the problem appears when considering the decision to stop. At a given time  $t$ , the DM compares the value of stopping now with that of continuing. Let  $\tilde{G}(t, \pi_t)$  be the value of the continuation option at time  $t$ , such that  $\tilde{G}(t, \pi_t) = \max\{\mathbb{E}[G(t+1, \pi_{t+1})|\pi_t]; \mathbb{E}[\tilde{G}(t+1, \pi_{t+1})|\pi_t]\}$ .

Following Kruse and Strack (2015), we use the concept of marginal incentive to sample in order to characterise the stopping time.

**Definition 1** (Marginal Incentive). *The marginal incentive,  $m : \mathbb{T} \times [0; 1] \rightarrow \mathbb{R}$ , measures the expected gain from proceeding to the next stage of the sampling process:*

$$m(t, \pi_t) = \tilde{G}(t, \pi_t) - c - G(t, \pi_t) \quad (2)$$

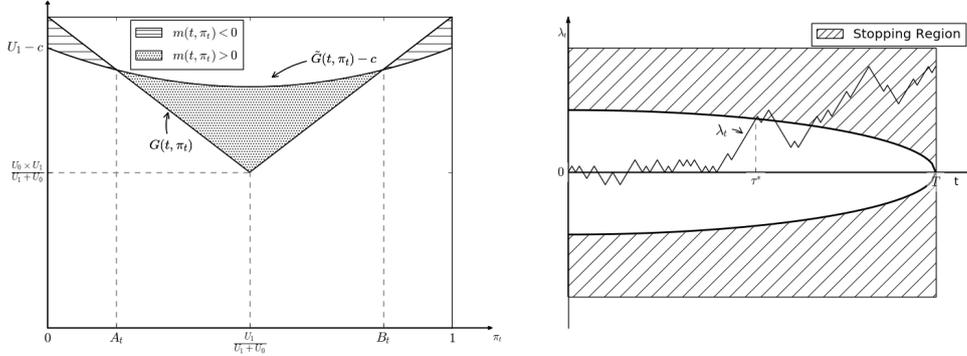
The optimal sequential sampling strategy is to continue sampling as long as there is a gain to it, that is whenever  $m(t, \pi_t) > 0$ . It can be shown that there are two levels of Bayesian beliefs that make the DM indifferent between sampling or not (*i.e.*  $m(t, \pi_t) = 0$ , see Tartakovsky et al. (2014)). We respectively call  $A_t$  and  $B_t$  the beliefs  $\pi_t$  solutions of equation (2). For these beliefs, the DM is indifferent between sampling and choosing  $d_t = a$  or  $d_t = b$  at time  $t$ . The left panel of Figure 1 summarises the decision rules to be followed by the DM.

The informational content of each signal can be summarised by a log-likelihood ratio,  $Z_t$ , which in the case of binary signals can only take two values ( $Z_t \in \{-\log(\frac{P(X_t=1|\theta=B)}{P(X_t=1|\theta=A)}); \log(\frac{P(X_t=1|\theta=B)}{P(X_t=1|\theta=A)})\}$ ). The rationale for using this measure is that it gives a simple additive form to the information that has been accumulated in the learning process:  $\lambda_t = \sum_{k=1}^t Z_k$ . It is then possible to define a stopping rule in the log-likelihood space, which summarises not only what decision the DM should make, but also when it should be made (Wald, 1945). Wald called this stopping rule a *sequential probability ratio test* (SPRT).

---

if  $d_t = b$ .

**Figure 1:** Theoretical representation of the sampling problem at a given time and in a dynamic setting.



*Note: Left panel:* Typical behaviour of the expected utility from stopping ( $G(t, \pi_t)$ ) and sampling ( $\tilde{G}(t, \pi_t)$ ), at a given stopping time. The dotted area represents the expected gains from sampling when beliefs are between  $A_t$  and  $B_t$ . The hatched area represents the expected losses from sampling when beliefs are outside the range  $[A_t, B_t]$ . **Right panel** typical behaviour of the SPRT statistics ( $\lambda_t$ ). The bold parabola represents the sampling frontier. The optimal strategy is to stop as soon as  $\lambda_t$  crosses this frontier (at  $t = \tau^*$ ).

**Definition 2** (Minimum Optimal Stopping Time). *The minimum optimal stopping time  $\tau^*$  is the first time  $\lambda_t$  hits a lower or an upper boundary, respectively  $a_t$  and  $b_t$ :*

$$\tau^* = \inf\{t \geq 0 : \lambda_t \notin [a_t, b_t]\} \quad (3)$$

Where  $a_t = \log \frac{A_t}{(1-A_t)}$  and  $b_t = \log \frac{B_t}{(1-B_t)}$  are the transforms of  $A_t$  and  $B_t$  in the log-odds space. The right panel of Figure 1 shows an example of path followed by  $\lambda_t$  over time, and the optimal stopping rule.

Wald and Wolfowitz (1948) show that for a given level of acceptance of statistical errors (type I and type II), the SPRT stopping time is optimal in the sense that it minimises the expected sample size. In our case, it means that the SPRT minimises the expected number of signals needed to reach a given level of risk.

## 2.2 Our experimental setting

We provide here a numerical example of the OSS model and its optimal solution, for the setting which we will use in our experiment. Let's consider a situation where the two states of the world are represented by urns with different proportions of white and black balls. A DM receives binary signals which are informative of the real state of the world: balls drawn from the urn. Suppose that both states of the world are a priori equally likely ( $\mathbb{P}(\theta = A) = \mathbb{P}(\theta = B) = 0.5$ ) and let's assume that white balls are “-1” signals, while black balls are “+1” signals. Suppose urn A contains 6 white balls and 4 black balls ( $\mathbb{P}(X_t = 1|\theta = A) = 0.4$ ) and urn B contains 4 white balls and 6 black balls ( $\mathbb{P}(X_t = 1|\theta = B) = 0.6$ ). The first step towards estimating the stopping region is to construct the Bayesian belief of the DM. Let  $n_B$  ( $n_A$ ) be the number of black (white) balls accumulated. Let's define, the *signal intensity* as the sum of accumulated signals,  $S_t = \sum_{i=1}^t X_i$ . A negative (positive) signal intensity indicates that state of the world A (B) is more likely. The Bayesian belief after  $t$  signals is given by:

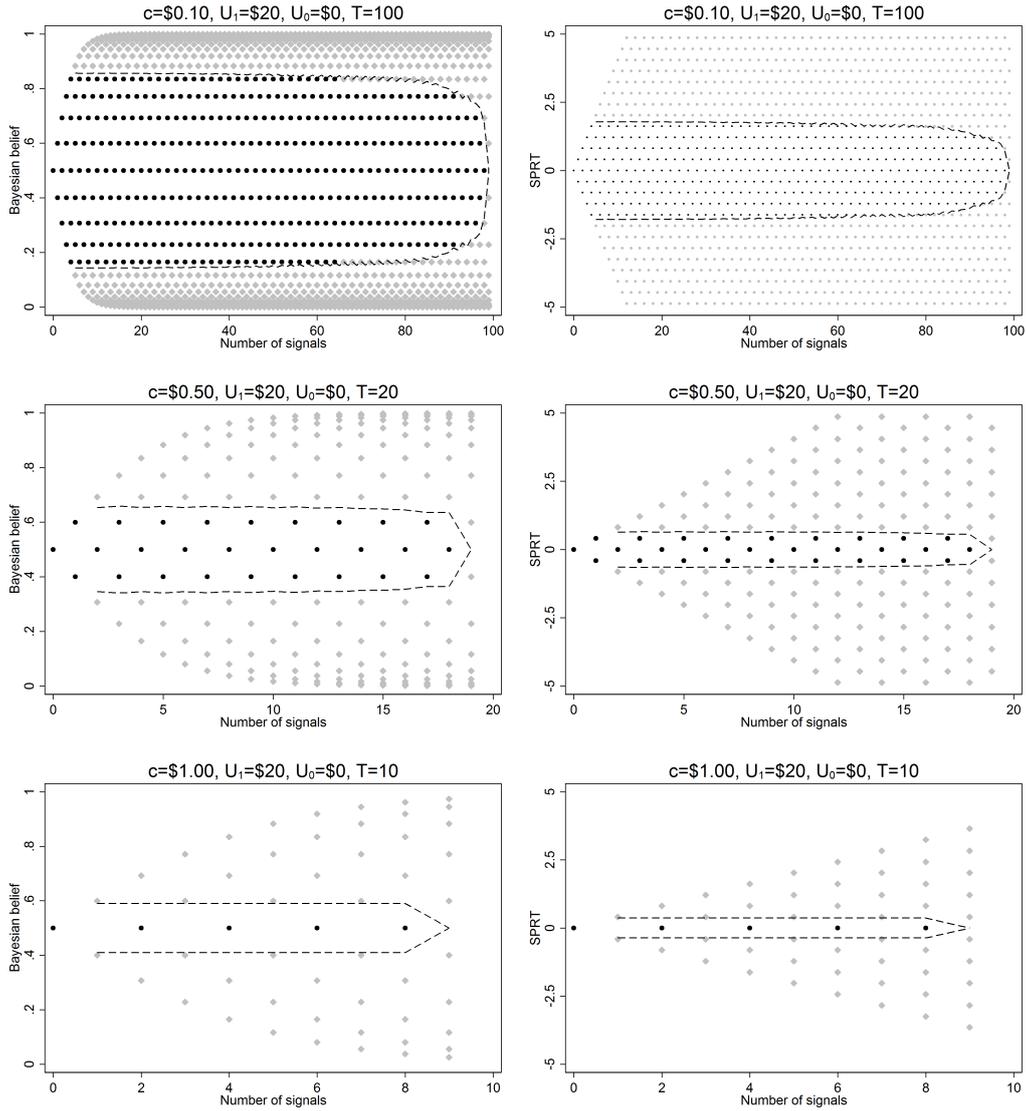
$$\pi_t = \frac{0.6^{n_B - n_W}}{0.6^{n_B - n_W} + 0.4^{n_B - n_W}}$$

It is possible to construct all the possible Bayesian beliefs after  $t$  draws, by looking at all the possible signals combinations. Let's consider a risk neutral DM with the following payoffs: being right yields \$20 ( $U_1 = 20$ ), being wrong yields \$0 ( $U_0 = 0$ ) and sampling an additional signal costs \$0.50 ( $c = 0.5$ ). For a given Bayesian belief ( $\pi_t$ ), it is possible to derive the expected payoff from stopping ( $G(t, \pi_t)$ ) in period  $t$ . In the last period ( $t = T$ ) the only possible choice is to stop, therefore the expected payoff from sampling at the period before the last one is  $\tilde{G}(t, \pi_T) = \mathbb{E}[G(T, \pi_T | \pi_{T-1})] - c$ . By backward induction, we can derive the expected payoff from sampling ( $\tilde{G}(t, \pi_t)$ ) at any period. From the expected payoffs of stopping and the expected payoffs of sampling, we can determine the boundaries  $A_t$  and  $B_t$ , that make the DM indifferent between sampling and stopping, by solving the equation  $m(t, \pi_t) = 0$ . Figure 2 represents these boundaries in terms of beliefs (left

panel) and in the corresponding log-likelihood space (right panel) for the three levels of cost used in our experiment.

When  $t$  is not too large, the optimal strategy is to stop when the difference between A and B signals crosses a threshold. For a low level of cost ( $c = 0.1$ ), the optimal strategy is to wait to have five more signals in one direction ( $\tau^* = \inf\{t \in ]5; 66[ \mid |S_t^*| = 5\}$ ). For  $c = 0.5$ , the DM should wait to have two signals in the same direction, that is  $\tau^* = \inf\{t \in \mathbb{T}; |S_t^*| = 2\}$ . Finally,  $c = 1$  is a degenerate case and the DM should gather only one signal. That is,  $\tau^* = 1$  and  $|S_1^*| = 1$ . As the experiment is done in a finite horizon, the optimal signal-strength threshold,  $|S_t^*|$ , is decreasing after a large number of draws, when  $t$  is getting close to  $T$ . This situation represents only a negligible part of the observed behaviour in our experimental data: respectively 0.38%, 0.37% and 1.19% of the observed stopping time for the low, medium and high cost conditions.

**Figure 2:** Plot of the sampling problem in the Bayesian beliefs space (left panel) and log-likelihood space (right panel) for  $c=0.10$ ,  $c=0.50$ ,  $c=1$ . Participants' endowment to buy signals is constant across the treatments, \$10. This endowment determine the maximum number of signals  $T$  which can be bought.



*Note:* The grey dots represent the stopping region. The black dots represent the sampling region. The prior is  $\mathbb{P}(\theta = A) = 0.50$ , a correct decision yields a payoff of  $U_1 = 20$  and a wrong decision a payoff of  $U_0 = 0$ . The dashed lines represent the estimated optimal frontiers.

## 2.3 Testable Hypotheses

This section summarises the features and predictions of the OSS model that can be tested in our experiment.

We define optimality as choosing to stop after the minimum optimal stopping time. As described above, we are able to characterise the optimal sampling strategy with a simple decision rule: a risk neutral agent should stop acquiring information when the log likelihood-ratio crosses one of the boundaries of the sampling region,  $a_t$  or  $b_t$ .

**Hypothesis 1. *Optimality.*** *An agent maximising expected payoffs follows the optimal sequential sampling strategy.*

The prediction of the OSS model regarding this change in incentives is unambiguous and intuitive. When the cost of information increases, the opportunity cost of going to the next round is higher. Hence, the higher the sampling cost, the lower the optimal number of draws.

**Hypothesis 2. *Reaction to incentives.*** *Individuals increase (decrease) their information acquisition when the cost of information decreases (increases).*

If **Hypothesis 1** is violated, we can expect participants to learn to make better decisions during the course of the experiment. Studies have shown that participants are able to learn to approximate optimal behaviour in demanding problems, such as equilibrium bidding in a double auction market (Friedman and Rust, 1993); or approximating the optimal timing of investment (Oprea et al., 2009). The optimal policy gives a clear stopping rule based on the difference between the number of white and black balls. Participants may be able to learn after some practice.

**Hypothesis 3. *Learning.*** *With experience individuals will learn to get closer to the optimal sequential sampling solution.*

According to the optimal sequential sampling model, framing the payoffs as gains or losses should not matter. However, there is extensive evidence that individual behaviour substantially changes depending on whether a prospect

is framed as leading to possible gains or losses (Kahneman and Tversky, 1979; Wakker, 2010). A simple way to introduce the idea that “losses loom larger than gains”, is to consider that when the DM faces a potential loss, the payoff  $V$  is subjectively transformed in  $\mu V$ , with  $\mu > 1$ . Mechanically, the higher stakes push the optimal frontier outwards and make the marginal incentive to sample greater.

**Hypothesis 4. *Framing Effect.*** *Framing the sampling problem as a way to avoid a loss increases the stopping time decision.*

## 3 Experimental Design

### 3.1 Design and Treatments

We design a sampling experiment where the participant has to decide what state of the world she is in, which determines her payoff.<sup>3</sup> The states of the world are represented by two urns (A and B), from which signals (balls) can be gathered at a unit cost. One of the urns contains 4 black balls and 6 white balls and the other one contains 6 black balls and 4 white balls. Both are depicted in Figure 12 in Appendix B. The participant has to guess which urn has been selected. She can pay a unit cost  $c$  to see a ball, which goes back into the urn after being observed. The balls are drawn automatically every 2s, and the participant decides when to stop the process. If she makes a correct guess, she earns \$20, and \$0 otherwise.<sup>4</sup> In each round, we gave participants a \$10 endowment to observe balls drawn from the urn.<sup>5</sup>

The participant takes part in 80 rounds. At the start of each round, one of the two urns is randomly selected by the computer with a probability  $\pi_0 = 0.5$ . One of the rounds is selected for payment at the end of the experiment. At all times during the experiment, all the relevant information

---

<sup>3</sup>This choice is identical to choosing an action  $d_t \in \{a, b\}$  mirroring the state of the world.

<sup>4</sup>All amounts are in Australian dollars.

<sup>5</sup>It is enough to reach the expected stopping time in each treatment. Time horizons were the following:  $T = 100$  in the \$0.10 condition,  $T = 20$  in the \$0.50 condition and  $T = 10$  in the \$1 condition.

is displayed on a screen (possible payoffs, number of draws, share of black and white balls in each urn, past draws). At the end of each round, participants receive feedback as to whether they chose the right urn and what is their payoff.

As we are able to precisely identify the optimal solution we can test whether participants approximate it (**Hypothesis 1**). Each participant enters one of three costs conditions (low \$0.10, medium \$0.50 and high \$1). Comparing decisions across these three treatments we can test whether participants react to the level of information cost (**Hypothesis 2**). By repeating the task multiple times and providing participants with feedback, we are able to test whether participants learn to approximate the optimal solution over time (**Hypothesis 3**). Finally, we frame the decision in two ways: payoffs are either presented as gains, or as losses relative to an initial endowment. This framing is neutral in terms of monetary incentives, the structure of payoffs being strictly identical across the two types of framing, and each participant enters only one of the two conditions.<sup>6</sup> These two types of framing allow us to test whether participants sample more to avoid losses than to secure gains (**Hypothesis 4**).

In summary, we have a  $2 \times 3$  between-subject design (Frame  $\times$  Cost) and the task is repeated 80 times.

## 3.2 Experimental Procedure

The experimental sessions were conducted between July 2015 and November 2016, in a large Australian university using z-Tree (Fischbacher, 2007). The participants were students from various faculties, and were recruited through ORSEE (Greiner, 2015). In total, 162 students took part in the experiment (with an average of 23.7 years old, and 56% males). Each session was approximately two hours long, including a short quiz testing participants' understanding of the game.

---

<sup>6</sup>In the loss frame we inform the participants that they are all given \$20 at the start of the session, 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.

After the 80 rounds were completed, the participants entered an incentivised risk preferences elicitation task (Holt and Laury, 2002). The subjects were then asked a few demographic questions and completed a cognitive reflection test (CRT, Frederick, 2005).<sup>7</sup> When all these tasks were completed, the participants were informed of their payoffs. The average payoff was \$28.5, including a \$5 show up fee and the risk preference elicitation procedure. One of the rounds the participants entered was randomly selected by the computer and the subjects were given the corresponding payoff.<sup>8</sup>

## 4 Results

### 4.1 Sampling behaviour

From these optimal rules, we can determine the corresponding average stopping times ( $\bar{\tau}^*$ ). To do so, we ran 100,000 simulations of a rational DM following the optimal decision rule for each level of information cost. From our simulations, on average 19.1 signals are required to reach a decision when  $c = 0.1$ , 3.84 when  $c = 0.5$  and 1 when  $c = 1$ . The theoretical signal intensities and corresponding average stopping times are compared with the participants' average decisions in Table 1. The table also presents the proportion of inconsistent decisions where participants choose an urn while their observed signals pointed to the other urn. Such choices represent only around 4% of all the choices and we do not take them into account in our analysis.

Overall participants make their decisions too quickly (sample too little information) when information is relatively cheap. Inversely, they hesitate too long (sample too much information) when information is relatively expensive. They stop after approximately 9 draws in the \$0.10 treatment, 7 draws in the \$0.50 treatment and 4 draws in the \$1 treatment. In the lower cost treatment this average is below the theoretical prediction, and in the

---

<sup>7</sup>The results of the CRT are provided in Appendix A.2.

<sup>8</sup>This has been shown to be an efficient way to allocate payment in repeated dynamic choice experiments, see Baltussen et al. (2012).

**Table 1:** Summary statistics from observed stopping times and signal intensity.

	\$0.1			\$0.50			\$1		
	N	Mean	sd	N	Mean	sd	N	Mean	sd
$\tau$ (Observed)	4560	9.42	8.82	4160	6.55	3.91	4240	4.19	2.48
$\tau$ (Theory)		19.1	13.83		3.84	2.64		1	0
$ S_\tau $ (Observed)	4560	2.85	2.03	4160	2.27	1.65	4240	1.74	1.37
$ S_\tau $ (Theory)		4.99	0.13		1.98	0.04		1	0
Inconsistent decisions	127	2.8%		190	4.6%		215	5.1%	

*Note:* 100,000 simulations were used to determine the average stopping time with a DM following the optimal stopping strategy. The last row reports the proportion of inconsistent decisions when participants made a choice in opposition with the direction of their signals.

two other costs treatments it is above it. The average stopping time is significantly different from the theoretical one in each treatment ( $p < 0.001$  for a Wilcoxon signed-rank tests, henceforth WSR, in each case). We use a between subject design and, in each treatment, individuals made 80 decisions. Looking at each individual, we reject the null hypothesis that the average number of draws is equal to the theoretical predictions for 142 out of 162 subjects (setting significance at  $p < 0.05$  for a WSR test).<sup>9</sup>

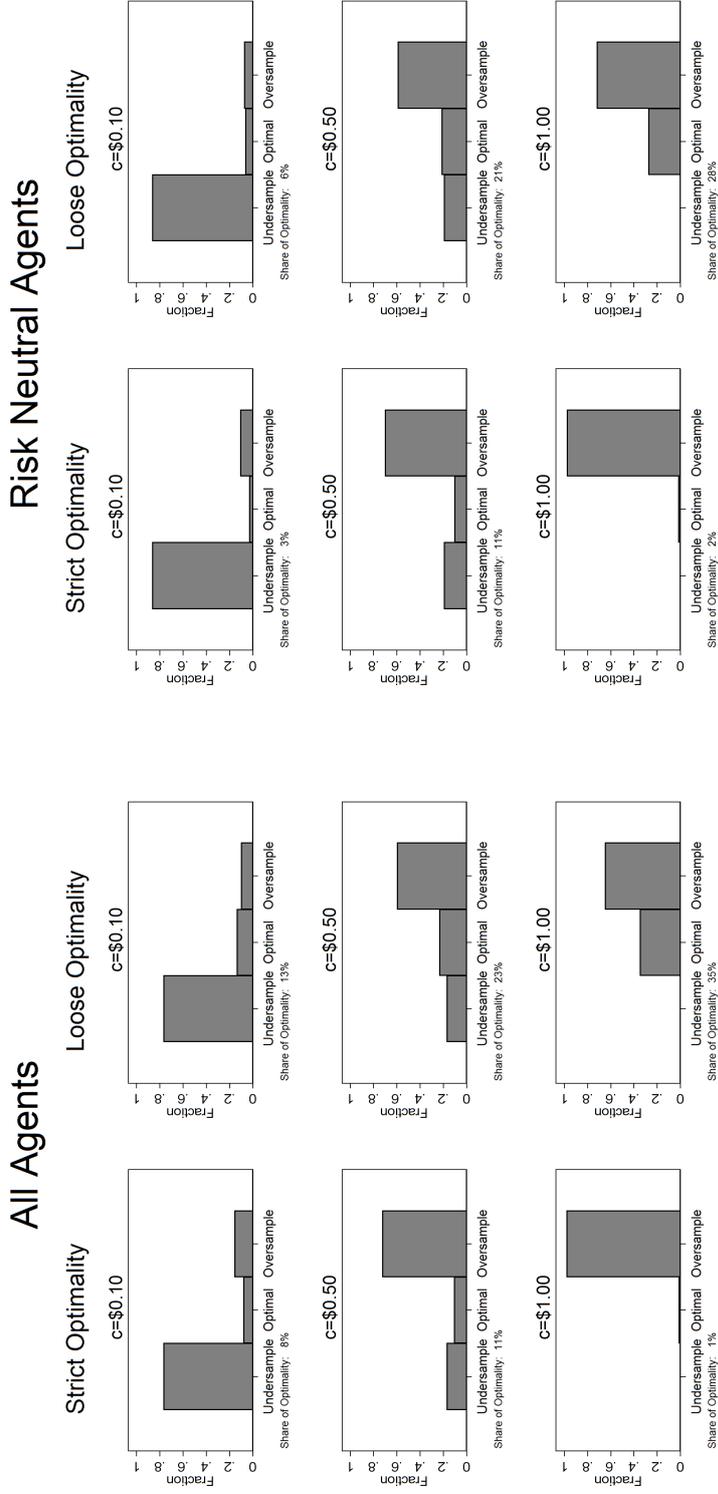
When the decision to stop has been taken, the signal intensity  $|S_\tau|$  reflects the level of uncertainty that was chosen by the DM. We observe the same deviation in this dimension, participants gathered too few signals in the \$0.10 treatment, and too many in the \$0.50 and \$1 ones. This information is provided in the lower part of Table 1. On average, in the \$0.10 treatment the participants had 2.85 signals in the same direction when they stopped - as opposed to 4.99 in theory. In the \$0.50 treatment, they had 2.27 signals in the same direction, while theory predicted 1.98 signals. In the \$1 treatment, the average signal intensity at the stopping time was 1.74 instead of 1. Looking at the individual level and comparing with the prediction from the condition the individual is in, the average signal density is significantly different the theoretical one for 133 of our 162 subjects (with a  $p < 0.05$  threshold for a WSR test). As a consequence, at the aggregate level, we strongly reject the hypothesis that the average signal intensity in the sample is equal to the

<sup>9</sup>All the WSR tests reported are clustered at the individual level and bootstrapped 10,000 times

theoretical predictions for each treatments ( $p < 0.001$  for WSR tests). As signal intensity is truncated at 0, medians may be a better measure of central location than means. Medians analyses confirm our previous results with median signal intensity statistically different from the theoretical predictions according to sign tests (median = 2.88,  $p < 0.001$  for \$0.10 treatment; median = 2.14,  $p = 0.0037$  for \$0.50 treatment; median = 1.71,  $p < 0.001$  for \$1 treatment)

The results of both the average stopping time and the average signal intensity suggest that participants stopped too early in the \$0.10 treatments, and too late in the \$0.50 and \$1 treatments. Figure 3 breaks down participants' choices between under/oversampling and optimal behaviour, defined as stopping at the right time. The first column of the figure shows this breakdown for the whole sample. The under/oversampling pattern is clear with the proportion of optimal behaviour staying low in each cost conditions. Over the three cost conditions, only 6.8% of decisions are in line with the optimal solution  $\tau^*$ .

**Figure 3:** Distribution of sampling behaviour by treatments, definition of optimality and risk preferences.



*Note:* Strict optimality is defined as stopping at  $\tau^*$  with a signal intensity  $S_{\tau^*}$ . Under and oversampling are respectively stopping before and after  $\tau^*$ . The loose optimality is defined as stopping between  $\tau^* - 1$  and  $\tau^* + 1$  with a signal intensity  $S_{\tau^*}$ .

One possibility could be that the optimal solution we discussed is for risk neutral DM while our participants may have different risk preferences. To make sure this is not what is driving the low levels of optimality, we use the risk preferences we measured using the Holt and Laury (2002) procedure to look at the participants who are risk neutral. The third column of Figure 3 displays the cumulative distribution of their choices. We do not observe that risk neutral participants perform better. Only 6.1% ( $N = 33$ ) of their decisions occurred at the optimal stopping time, whereas this figure is respectively 6.3% and 8.3% for risk averse and risk loving participants ( $N = 98$  and  $N = 31$ ). A further test of the impact of risk preferences on sampling behaviour is provided in Section 5.1.1.

Looking only at the number of times participants precisely stopped at the optimal threshold of information is quite a stringent test of optimality. It may give the impression that participants fail to optimise even though they may be approximating the optimal strategy quite well. To address this concern, we defined a “loose” optimal behaviour such that being one draw away from the right stopping time  $\tau^*$  is defined as optimal.<sup>10</sup> The second and fourth columns of Figure 3 show the corresponding breakdown of participants’ choices over the whole sample and for risk neutral participants. The rate of loose optimality is 23.5% in the whole sample and 20% for risk neutral agents. For all participants, this figure ranges from 13% in the \$0.10 treatment, 23.5% in the \$0.50 and 34.8% in the \$1 treatment. The sampling pattern persists in the “loose” definition of optimality, as we still observe 77% of undersampling for  $c = 0.1$ , and respectively 59% and 65% of oversampling for  $c = 0.5$  and for  $c = 1$ .

The cost of the deviation we observe is substantial. In the \$0.10 treatment, the average expected payoff is \$2 lower than the optimal level (\$15.65 vs \$13.61). In the \$0.50 treatment the average payoff is \$1 lower than model prediction (\$11.42 vs \$10.15). In the \$1 treatment, participants earn \$2.5 less than expected (\$10.47 vs \$7.96). These deviations correspond to foregone payoffs worth between 10% and 25% of potential payoffs.

---

<sup>10</sup>In the \$0.1, \$0.50, and \$1 treatments, this is equivalent to increasing the size of the sampling set by 10%, 25% and 100% respectively.

**Result 1.** *Individuals deviate systematically from the optimal solution.*

- (i) For **low** information costs, participants **undersample** relative to the optimal solution.
- (ii) For **high** information costs, participants **oversample** relative to the optimal solution.
- (iii) The same pattern is observed among the sub-sample of participants who are risk neutral.

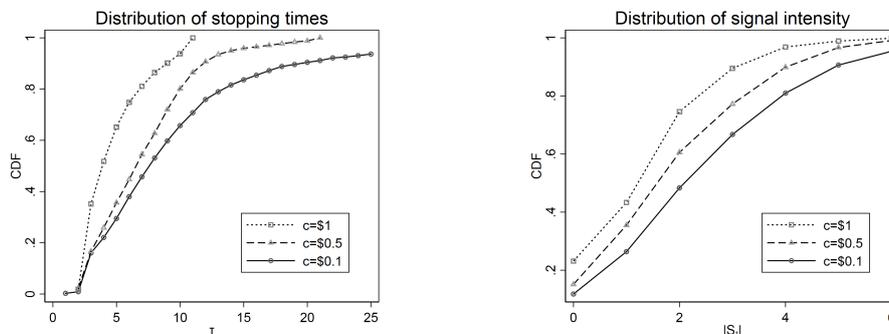
## 4.2 Reaction to incentives

A possible explanation to these results is that participants did not react to the change in incentives. It could be that they always chose the same signal intensity, regardless of the cost incurred to sample. Looking at the distribution of stopping times and signal intensities for the three levels of costs it appears that, on the contrary, participants reacted to incentives. The right panels of Figure 4 represents the cumulative distribution functions (CDF) of stopping times and signal intensities for each level of sampling cost. In each case, we observe a clear first stochastic dominance from the distributions with lower costs relative to distributions with higher costs. Kolmogorov-Smirnov tests (clustered at the individual level and bootstrapped 10,000 times) confirm that distribution for each level of cost significantly differs from other ones both for stopping times and for signal intensity ( $p < 0.001$  in each cases except for the test of signal intensity between high and medium cost -  $p = 0.003$ ). These results show that participants have reacted to the changes in the cost of information as predicted.

**Result 2.** *Participants react to changes in incentives by reducing information acquisition for higher costs.*

This result supports our **Hypothesis 2**: more costly information leads to earlier stopping times.

**Figure 4:** Cumulative distribution of stopping times and signal intensity.



*Note:* For clarity purpose we restrict the figures to stopping times below 26 and signal intensity below 6. The complete figures can be found in Appendix.

### 4.3 Learning

One of the main findings in Oprea et al. (2009) is that after a learning phase, participants were able to approximate the optimal behaviour in the stopping problem. In our case, we are able to locate how far a stopping decision is from the average minimum optimal stopping time ( $\tau - \bar{\tau}^*$ ), and the difference between the resulting signal intensity and the average optimal one ( $|\mathcal{S}_\tau| - |\bar{\mathcal{S}}_{\tau^*}|$ ).

Figure 5 depicts the evolution of the participants' average stopping times  $\tau$  and the resulting signal intensities  $|\mathcal{S}_\tau|$ , throughout the 80 rounds of the experiment. We observe that the gap between observed behaviour and the optimal solution decreases over time for the two treatments where information acquisition is relatively costly ( $c = 0.5$  and  $c = 1$ ). These trends are statistically significant for the average stopping time ( $p = 0.03$  for  $c = 0.5$  and  $p < 0.001$  for  $c = 1$ , ordinary least squares, henceforth OLS <sup>11</sup>) and for the average signal intensity (respectively  $p = 0.1$  and  $p < 0.01$ , OLS). If we analyse the last 20 periods of both treatments, we reject the hypothesis that the distance to the optimal stopping time, as well as the distance to the optimal number of draws, is equal to zero ( $p < 0.001$  for WSR tests in all cases), meaning that oversampling persisted. The average stopping time just

<sup>11</sup>All regressions are clustered at the individual level.

reaches the boundary of the loose optimality at the end of the experiment in the \$1 condition.

These results suggest that participants learned to make better decisions over time in the conditions with  $c = \$0.50$  and  $c = \$1$ . In the treatment where information was the most costly,  $c = \$1$ , participants were able to approximate the optimal solution over time. In contrast, we do not find any convergence towards the optimal solution in the \$0.10 treatment.<sup>12</sup>

These results support **Hypothesis 3** when sampling is relatively costly. We purposely chose to have a large number of repetitions of the decision situation to give enough time for learning to take place. Our findings suggest that learning takes place, except when the sampling is relatively cheap.

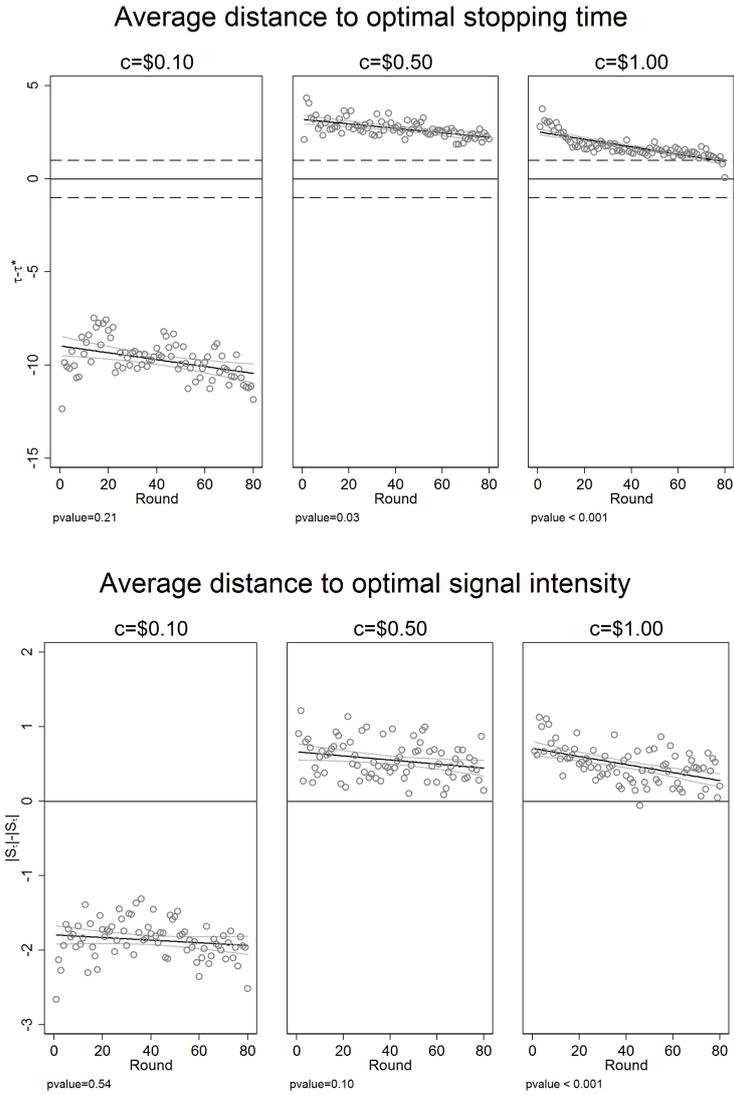
**Result 3.** *We observe learning and a convergence towards the optimal sequential sampling solution when sampling is relatively costly.*

- (i) Participants tend to learn and get closer to the optimal strategy when information acquisition is costly.*
- (ii) No learning seems to take place in the treatment where information acquisition is relatively cheap. The undersampling pattern does not decrease.*
- (iii) In our treatment with the highest cost of information acquisition, participants learn to closely approximate the optimal strategy.*

---

<sup>12</sup>The slope of the regression line is neither statistically significant for the average stopping time ( $p = 0.21$ ) nor for the signal intensity ( $p = 0.54$ ).

**Figure 5:** Evolution of the distance to the optimal stopping time and signal intensity.



*Note:* The top panel shows the average distance to the minimum optimal stopping time ( $\tau - \bar{\tau}^*$ ) for each round. The dashed lines represent the “loose” optimality condition, such that stopping one ball away from the optimal rule is regarded as optimal. The bottom panel shows the average distance to the optimal signal intensity resulting from the stopping time ( $|S_\tau| - |\bar{S}_{\tau^*}|$ ). The black line represents the linear fit and the outer grey lines are the 95% confidence intervals. The value displayed below each graph is the p-value used in testing the null hypothesis that the OLS regression coefficient is 0 (clustered at the individual level).

## 4.4 Framing Effects

In the experiment we explicitly framed the decisions as either a way to make a gain or to avoid a loss, in order to test for an effect of loss aversion. Our hypothesis is that loss aversion raises the subjective value that is at stake, hence increasing the sampling region. We find that it is not the case.

Table 2 compares the average stopping times and signal intensities for the gain and loss framing for each level of cost. We do not observe significant differences between the two types of framing. The differences in sampling behaviour are small and not significant. Overall, there is no evidence that participants sample more in the loss frame.

**Result 4.** *Participants did not significantly accumulate more information in the loss condition.*

To summarise, we find that participants substantially deviate from optimality in a systematic way: they oversample information when the sampling cost is high, and undersample it when the cost is low. However, participants respond to incentive changes and decrease the amount of information collected when the cost increases. The overall performance improves as time passes, without fully converging towards the optimal solution. In addition, we find that framing the problem as a way to avoid a loss does not increase sampling.

**Table 2:** Mean number of draws ( $\bar{\tau}_{Gain}$  and  $\bar{\tau}_{Loss}$ ) and signal intensity ( $|\bar{S}_\tau|_{Gain}$  and  $|\bar{S}_\tau|_{Loss}$ ), by frame.

	\$0.10	\$0.50	\$1
$\bar{\tau}_{Gain}$ vs $\bar{\tau}_{Loss}$	9.24 vs 9.49	6.34 vs 6.78	4.56 vs 3.81
Mann-Whitney ( $p$ )	0.35	0.42	0.16
$ \bar{S}_\tau _{Gain}$ vs $ \bar{S}_\tau _{Loss}$	2.83 vs 2.89	2.17 vs 2.36	1.83 vs 1.64
Mann-Whitney ( $p$ )	0.60	0.24	0.26

*Note:* Mann-Whitney tests bootstrapped with 10,000 replications and clustered at the participant level.

## 5 Potential Explanations

In this section we discuss potential explanations to the pattern of deviations from optimality observed in our experiment: undersampling when the cost of information is low, and oversampling when the cost of information is high.

The OSS model rests on three main assumptions. First that DMs are risk neutral utility maximisers; second that they react to information and update their beliefs accordingly; third that they are forward looking and able to compute the expected value of sampling. Based on findings from previous studies, we explore alternative assumptions that could account for the observed behaviour in our experiment.

### 5.1 Wider class of utility functions

In Section 4 we ruled out that risk neutral participants behaved optimally. As the assumption of risk neutrality is quite restrictive, we consider here different types of utility functions that are relevant in a sequential sampling framework.

#### 5.1.1 Risk aversion

When considering risk aversion, an intuition could be that risk averse DMs are unwilling to take a risky gamble and prefer to sample to reduce uncertainty. However, the link between risk aversion and the optimal amount of information to collect is, in theory, ambiguous. It is *a priori* not clear whether an increase in risk aversion systematically leads to an increase in sampling (see Radner and Stiglitz (1984); Treich (1997); Gollier (2004) for a description of this problem). When information is costless and only a few signals can be accumulated, an increase in risk aversion can be associated with an increase or decrease in sampling, depending on the starting level of risk aversion. The case when information is costly, to the best of our knowledge, is yet to be investigated.

While theory does not provide general predictions, in our setting we can investigate the role of risk aversion by simulation. To do so, we solve the

optimal stopping problem by backward induction with different levels of risk aversion. We consider DMs with a CRRA utility function (Wakker, 2010), such that the utility of making the right guess is:

$$U_1 = \begin{cases} \ln(20) & \text{if } \gamma = 1 \\ \frac{20^{1-\gamma}}{1-\gamma} & \text{if } \gamma \neq 1 \end{cases}$$

Figure 6 shows the average optimal stopping time from these simulations for a wide range of risk aversion levels. Risk aversion does not have a monotonic effect on the sampling strategy. Starting from a situation where the DM is risk loving ( $\gamma < 0$ ), an increase in risk aversion will first leads the DM to sample less, then to sample more.

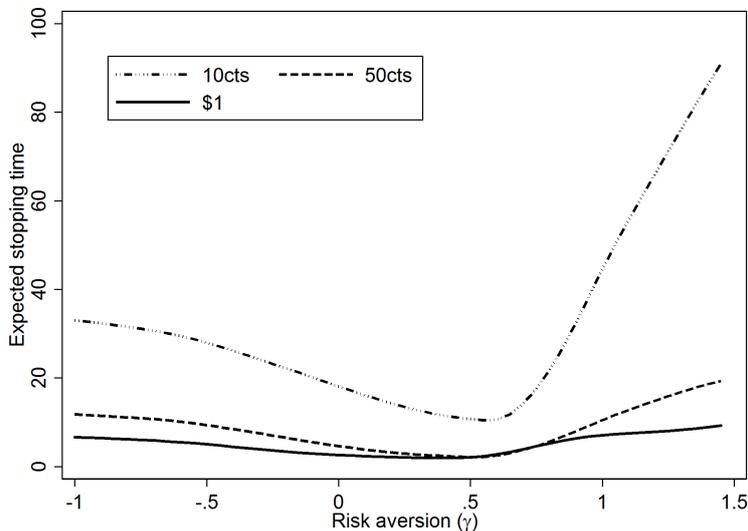
Could risk aversion explain our results? On one hand, it could explain the undersampling results by participants having  $\gamma$  parameters between 0.3 and 0.9, which would be regarded being as between mildly and very risk averse - not far from what is typically observed in experiments. On the other hand, the oversampling result would require participants to be either very risk averse, or very risk loving in the \$0.50 and \$1 treatments.

It is unlikely that, by chance, randomly allocated participants would have very different risk aversion levels across our different treatments. Indeed, using the measure of risk aversion from the Holt and Laury (2002) procedure, we find no difference in average risk aversion between participants from different cost conditions ( $p = 0.49$  for a Kruskal-Wallis test).

To further assess how risk preferences affect sampling behaviour, we estimate for each participant the optimal stopping time given their elicited  $\gamma$  parameter using the Holt and Laury (2002) procedure. The results of this exercise are displayed in Appendix A.3. We find little evidence that risk aversion explains our results.

To sum up, while risk aversion is an important parameter in the decision to sample or not, there is no evidence that it would be the main driving force behind our results.

**Figure 6:** Risk aversion and average optimal stopping times



*Note:* This figure represents the expected stopping time for each of our treatments, as a function of the risk aversion parameter in the CRRA utility function.

### 5.1.2 Loss aversion

When investigating the effect of framing on sampling decisions, we assumed that framing would lead the participants to perceive the stakes as larger in the loss frame. However, it could also be that they view the sampling cost as a loss itself, thereby affecting sampling decisions. We conducted simulations to test whether this possibility is correct and analyse more the impact of loss aversion on payoffs.

We consider a standard loss aversion parameter ( $\mu = 2$ , see Abdellaoui et al. (2007) for a review), and study its impact on decisions in each of our framing condition.

In the gain frame, if loss aversion affects the sampling costs, then each draw is perceived as costing  $\mu \times c$ . All else being equal, higher subjective sampling costs reduce the optimal stopping time. As a consequence, we find that loss aversion leads theoretical predictions to be closer to observed choices in the \$0.10 treatment. But on the contrary, theoretical predictions get even further from observed choices in the treatments \$0.50 and \$1 where

participants sampled more signals than predicted by theory. Adding loss aversion on the sampling cost makes this oversampling worse.

In the loss frame, assuming that the framing triggers loss aversion, if sampling costs are affected by loss aversion, then both the costs of sampling  $c$  and the payoffs  $U_0, U_1$  are subjectively multiplied by  $\mu$ . As the optimal solution does not depend on the scale of payoffs, it is the same as in a gain frame where neither payoffs nor costs are affected by loss aversion. We observe therefore the same pattern of under and oversampling as in the gain frame without loss aversion.

As a consequence, loss aversion on sampling costs does not provide an explanation to the pattern of under and oversampling in the participants' choices.<sup>13</sup>

### 5.1.3 Regret

In a recent study, Viefers and Strack (2014) find that regret plays a major role in an optimal stopping experiment. Regret, as defined by Loomes and Sugden (1986), captures the idea that agents are affected by the counterfactuals of their action (Bleichrodt and Wakker, 2015). In a dynamic setting a DM can not only feel regret over past decisions, but also anticipate the possible regret over future decisions.

First, regret over past decisions implies that DMs are reluctant to stop after having had the opportunity to stop at a point where the expected value of success was high. The optimal strategy for a regret averse DM that has missed the *ex ante* optimal strategy is to wait until the process gets back to this level, leading to longer stopping times. Second, anticipated regret arises because DMs are aware that they might experience regret and account for it in their decision making. Viefers and Strack (2014) show that if the effect of anticipated regret is as strong as regret over past decisions, stopping times are not affected by its introduction. We therefore consider a simple case where regret exists without anticipated regret.

---

<sup>13</sup>Table 6 in Appendix A.5 displays the theoretical predictions in each framing, whether costs are considered or not as losses.

Following Viefers and Strack (2014), we model regret over past decisions as discounting the value of stopping by the maximum value the DM could have received ( $G_{MAX}$ ) multiplied by a factor,  $\kappa \in [0; 1[$ , which represents the intensity of regret. Hence the marginal incentive to sample with regret is:

$$m^R(t, \pi_t) = \tilde{G}(t, \pi_t) - c - [G(t, \pi_t) - \kappa \times G_{MAX}]$$

As  $\kappa$  and  $G_{MAX}$  are positive numbers, introducing regret mechanically leads to a higher marginal incentive to sample. Hence it might be a potential explanation to the oversampling pattern we observe. We model our regret averse DM as risk neutral.<sup>14</sup>

We estimate the regret parameter  $\kappa$  by maximum likelihood using a random utility approach where the probability to sample at a given period in the experiment is:

$$\mathbb{P}(\text{Sampling}) = \mathbb{P}(m^R(t, \pi_t) > c) = \Phi(m^R(t, \pi_t) - c)$$

Where  $\Phi$  is the CDF of a standard normal distribution. We estimate this model, at the individual level, using the following log-likelihood function:

$$\ln L_i = \sum_{n_i=1}^{N_i} Y_{n_i} \times \ln [\mathbb{P}(\text{Sampling})] + (1 - Y_{n_i}) \times \ln [1 - \mathbb{P}(\text{Sampling})]$$

Where  $Y_{n_i} = 1$  if the decision to sample has been taken, and 0 otherwise, and  $N_i$  is the total number of decision that was taken by individual  $i$ .

The results of this estimation are displayed in Table 3. We find an average regret coefficient,  $\kappa$ , of respectively 0.01, 0.06 and 0.09 in the \$0.1, \$0.50 and \$1 treatments. We find that this regret coefficient is significantly different from zero for 150 out of our 162 subjects (setting the significance threshold at  $p = 0.05$ ).

However, we strongly reject the hypothesis that these parameters are

---

<sup>14</sup>Jointly estimating regret and risk aversion leads to non trivial tractability issues as one is estimated directly based on past decisions while the other one is introduced by backward induction.

similar across treatments ( $p < 0.01$  for pairwise Mann-Whitney tests). If the pattern of under and oversampling was explained by regret, one would expect a given level of regret to explain choices across treatments. Instead, the fact that the regret parameter varies substantially suggests that it may act as an adjustment parameter helping the model to fit the data better in different situations. This regret parameter varying between treatments might be capturing unexplained variations by the model, rather than regret *per se*. As a consequence, it is not clear that regret provides a unified explanation for the observed pattern.

**Table 3:** Mean and standard deviation of the individual estimates we derived for the regret coefficient ( $\kappa$ ).

	c=0.1	c=0.50	c=1
Mean	0.01	0.06	0.09
sd	0.02	0.03	0.03

## 5.2 Misuse of information

While Bayesian updating is a cornerstone of our model, it is well documented that participants often violate Bayes’ rule (Kahneman and Tversky, 1979; El-Gamal and Grether, 1995). The problem we are considering is particularly prone to two types of biases. First it is possible that participants use a “rule of thumb” (Kahneman, 2003), whereby they take a constant number of signals and do not adjust their behaviour to new information. Second participants could exhibit a confirmation bias (Rabin and Schrag, 1999), whereby information is interpreted as confirming one’s own one belief.

### 5.2.1 Rule of thumbs

An explanation to our results could be that participants do not react to new information and use “rule of thumbs” (e.g. always selecting the same number of draws). A first glance reveals that no participant chose the same stopping time throughout the experiment.

To further assess this possibility we test whether the decision in trial  $t - 1$  on the number of draws ( $n_{t-1}$ ) predicts the number of draws in the next trial ( $n_t$ ). To do so, we estimate a first-order autoregressive Poisson model (Martin et al., 2012):

$$n_t = \rho \circ n_{t-1} + \varepsilon_t$$

Where:

$$\rho \circ n_{t-1} = \sum_{s=1}^{n_{t-1}} e_{s,t-1}$$

$$e_{s,t-1} = \begin{cases} 1 & \text{with probability } \rho \\ 0 & \text{with probability } 1 - \rho \end{cases}$$

$$\varepsilon_t \sim i.i.d. \mathbb{P}(\lambda)$$

The notation  $\circ$  is the binomial thinning operator, which is used to ensure that the integer nature of data is preserved.<sup>15</sup>  $P(\lambda)$  stands for the standard Poisson process with parameter  $\lambda$ . As  $\rho \circ n_{t-1}$  is the sum of  $n_{t-1}$  *i.i.d.* random variable, we get  $\mathbb{E}[n_t | n_{t-1}] = \rho \times n_{t-1} + \lambda$ . The autocorrelation parameter  $\rho$  can be estimated by maximum likelihood at the individual level (Weiß, 2008).

The average autocorrelation parameter is  $\bar{\rho} = 0.47$ . It is significantly not different from zero in 75% of cases (setting the significance threshold at  $p = 0.05$ ). Therefore, for most participants the number of draws in the past round does *not* significantly predict the current round decision. It suggests that participants did not use the same stopping rule throughout the rounds and did adapt their sampling behaviour to the signal intensity they observed. As a consequence, the pattern of under and oversampling does not seem to be driven by the use of simple rule of thumbs.

### 5.2.2 Confirmation bias

A standard finding in Bayesian updating tasks is that participants exhibit a confirmation bias, whereby they tend to interpret information in a man-

---

<sup>15</sup>It sums  $n_{t-1}$  Bernoulli variables ( $e_{s,t-1} \sim \text{Bern}(\rho)$ ).

ner biased towards confirming their existing beliefs. Experimental studies have shown that participants seek more confirmatory evidence when evaluating competing hypotheses (Jones and Sugden, 2001) or place more weight on confirming evidence than disconfirming ones (Charness and Dave, 2017). From a theoretical stance, this bias has been theorised by Rabin and Schrag (1999) and potentially arises from a preference for consistency (Yariv, 2002) or bounded memory (Wilson, 2014).

Following Rabin and Schrag (1999), one way to introduce confirmation bias in our framework is to consider that the DM perceives signals differently than what they actually are. Denote  $\tilde{n}_B$  ( $\tilde{n}_A$ ) the perceived number of black (white) balls that have been drawn. Let's assume that the distortion takes a simple form, such that the DM linearly overestimates the value of the draws that confirm her prior. That is she thinks that each black ball is worth  $\alpha > 1$  white balls, if her prior is greater than 0.5. If the DM has an equal number of black and white balls, there is no bias. It means that the belief at a given period  $t$  is:

$$\tilde{\pi}_t = \begin{cases} \frac{0.6^{n_B - n_W}}{0.6^{n_B - n_W} + 0.4^{n_B - n_W}} & \text{if } \pi_{t-1} = 0.5 \\ \frac{0.6^{\alpha n_B - n_W}}{0.6^{\alpha n_B - n_W} + 0.4^{\alpha n_B - n_W}} & \text{if } \pi_{t-1} > 0.5 \\ \frac{0.6^{n_B - \alpha n_W}}{0.6^{n_B - \alpha n_W} + 0.4^{n_B - \alpha n_W}} & \text{if } \pi_{t-1} < 0.5 \end{cases}$$

To test this hypothesis, we conducted simulations for various levels of  $\alpha$  ( $\alpha \in [1; 3]$ , with increment of 0.025). For each level of bias, we computed the sampling and stopping payoffs. Then for each sequences of draws we looked at the minimum optimal stopping time based on biased payoffs. At an individual level, the estimated bias parameter is the one that produces the best fit between observed and predicted stopping times. Table 4 summarises the outcome of our simulations. In the \$0.10 treatment we find average confirmation bias  $\alpha$  of 1.91 and a median of 1.96. This bias explains the decisions of participants, with 17% of them making the “optimal” decision

given these biased beliefs and 61% of participant stopping within one draw of this best choice. In the \$0.50 treatment, we find an average confirmation bias  $\alpha$  of 1.71, with a median of 1.4. This model with confirmation bias explains the decisions of participants in 16% (strict definition) and 72% (loose definition) of cases. In the \$1 treatment, the mean confirmation bias  $\alpha$  is 1.80 and the median is 1.28. Conditional on these biased beliefs, the decisions are “optimal” respectively in 24% (strict definition) and 91% (loose definition) of cases.

These results suggest that the confirmation bias is a possible explanation for our results. A majority of participants make decisions within one ball of the (subjectively) best strategy (i.e. best given the beliefs of the DM, even if wrong). To gain intuition as to how the confirmation bias affects decision making, we compare how a Bayesian and biased DM would update their beliefs when receiving a sequence of signals in the same direction. Figure 7 displays beliefs, and the marginal contribution of the last draw to beliefs, for a Bayesian DM and a DM with confirmation bias. For the first piece of information, a biased DM and a Bayesian DM update their belief in the same way. Then, for each new piece of confirming information the biased DM updates her belief faster than the Bayesian DM. However, after a certain point, as the value of information is bounded, confirming draws contribute less in updating in the biased case than in the Bayesian case.

This simulation helps make sense of how confirmation bias can explain part of the undersampling in the low cost treatment and oversampling in the high costs treatments. In the \$0.50 and \$1 treatments, where the number of draws is typically low, the biased DM overestimates the marginal value of a draw compared to a Bayesian DM. On the other hand, in the \$0.10 treatment, where the number of draws is typically high, the biased DM underestimates the value of a draw, compared to a Bayesian DM.

This intuition is confirmed by looking at the marginal incentive to sample ( $m(t, \pi_t)$ ).<sup>16</sup> In the \$0.10 treatment, when considering to draw the optimal signal intensity (i.e. going from  $S_\tau = 4$  to  $S_\tau = 5$ ), the marginal incentive to

---

<sup>16</sup>All the computations that follow are conducted using the average estimated parameters from Table 4

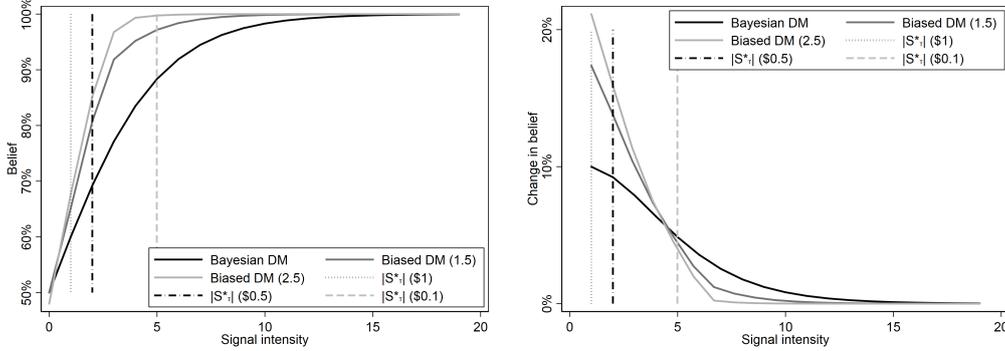
**Table 4:** Summary of the estimation of individual confirmation bias parameters.

	$\alpha$		Optimality rate	
	mean	median	strict	loose
\$0.10	1.91	1.96	17%	61%
\$0.50	1.71	1.40	16%	72%
\$1	1.80	1.28	24%	91%

sample is slightly above zero in the Bayesian case (0.06), compared to -0.21 in the biased case. In the \$0.50 treatment, the average incentive to sample at the theoretical expected stopping time ( $\bar{\tau}^* = 4$ ) is -0.15 in the Bayesian case, compared to 1.72 in the confirmation bias one. In the \$1 treatment the same observation holds, as the incentive to sample is -0.58 in the Bayesian case, compared to 1.77 in the biased case when  $\bar{\tau}^* = 2$ . As a consequence a biased DM will stop before a Bayesian DM in the \$0.10 treatment and after a Bayesian DM in the treatments \$0.50 and \$1.

Moreover, and contrary to the regret model, the estimated parameters are of similar magnitude in all treatments. If we compare the distribution of the parameters, we cannot reject the hypothesis that they are the same across treatments ( $p = 0.56$  for a Kruskal-Wallis test).

**Figure 7:** Belief updating for a Bayesian and biased DM



*Note:* The figures represent how subjective beliefs are affected by a confirmation bias ( $\alpha = 1.5$  and  $\alpha = 2.5$ ). **Left panel:** How beliefs change as a function of signal intensity for a Bayesian and a biased DM. **Right panel:** Marginal effect on subjective beliefs of an increase in signal intensity (the curves are the derivatives of the left panel curves). In both panels, vertical lines represent the signal intensity at the stopping time which is optimal (conditional on the subjective beliefs of the DM).

### 5.3 Myopic decision making

The OSS model assumes that DMs are forward looking and able to compute the sampling expected payoff (possibly by backward induction). In contrast with this hypothesis, Gabaix et al. (2006) find that participants in a search experiment are partially myopic. More specifically, at each stage of the search problem they behave *as if* the next period was the last one (and reiterate naively at each period). In our framework, it means the expected value from sampling would only be the expected value from stopping in the next period ( $\tilde{G}(t, \pi_t) = \mathbb{E}[G(t+1, \pi_{t+1})|\pi_t]$ ).

To test this hypothesis we computed the expected stopping time and signal intensities for each of the treatments. The theoretical prediction of this model is similar across treatments. Given our parameters, all agents should only acquire one signal. This eventuality is at odds with the typical number of signals observed by our participants. Moreover, we do observe that participants behave markedly differently in the three treatments, in contrast with the myopic model's prediction. As a consequence, the myopic model does not seem to be an explanation for the pattern of over and undersampling we observed.

## 5.4 Overview of explanations

We investigate different possible explanations for the deviations from optimality we observe. The undersampling can be partially explained by risk aversion or myopic decision making. But these explanations cannot help understand the existence of oversampling in the treatments where the information was most expensive. Oversampling may be partly explained by regret, but the variations of the regret parameter across the different treatments cast doubt on regret as a unifying explanation of the observed behaviour. The only explanation which we find to possibly reconcile both over and undersampling is the existence of a confirmation bias.

## 6 Conclusion

This study has investigated to what extent people are able to optimally hesitate between two options with uncertain values. To do so we have designed an experiment where participants have imperfect information about the payoffs associated with two possible choices, A and B. These choices can be thought as representing goods to purchase, investments to make or diagnostic decisions to consider. Participants can choose to wait and learn more information about the values associated with these choices, or to stop (at any moment) and select one of them. In our experimental setting, we know exactly the participant's information and the alternatives' payoffs. We are therefore able to determine the optimal behaviour in this situation: the optimal sequential sampling strategy. We can then measure whether (and how) participants deviate from optimality.

Our results can be summarised as follows. First, we find that participants deviate substantially from the optimal strategy. When sampling is relatively expensive, participants *waited too long* to collect information before making a decision; on the other hand, when it is relatively cheap, they *did not wait enough* and made their decision too early. Second, we find that participants learn over time to improve how long they wait before making a choice when information is relatively costly. In the treatments with costly information, the

oversampling of information decreases over the duration of the experiment. At the end of the experiment the average behaviour of participants is close to the optimal solution. However, when information is cheap we do not observe improvement over time with participants consistently collecting less information than they should under the optimal sampling strategy.

We test for a wide range of theories that could explain our results and find that, among the possible explanations, the confirmation bias is able to explain most of the observed under/oversampling. When introducing a confirmation bias, approximately 75% of the decisions are (loosely) the best decisions the DM should make given her biased beliefs.

The economic situations we investigate here are pervasive: in a wide range of contexts, decision makers face a choice between making a decision now or waiting to know more about the different options. Our results suggest that the optimal sequential sampling framework can be a good benchmark to model the behaviour of experienced decision makers. However, our results also open new questions as to how and why people deviate in systematic ways from the optimal solution.

## References

- ABDELLAOUI, M., H. BLEICHRODT, AND C. PARASCHIV (2007): “Loss aversion under prospect theory: A parameter-free measurement,” *Management Science*, 53, 1659–1674.
- BALTUSSEN, G., G. T. POST, M. J. VAN DEN ASSEM, AND P. P. WAKKER (2012): “Random Incentive Systems in a Dynamic Choice Experiment,” *Experimental Economics*, 15, 418–443.
- BLEICHRODT, H. AND P. P. WAKKER (2015): “Regret theory: A bold alternative to the alternatives,” *The Economic Journal*, 125, 493–532.
- 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, 700.
- CAPLIN, A. AND D. MARTIN (2015): “The dual-process drift diffusion model: evidence from response times,” *Economic Inquiry*.
- CHARNESS, G. AND C. DAVE (2017): “Confirmation bias with motivated beliefs,” *Games and Economic Behavior*, 104, 1–23.
- DELLA SETA, M., S. GRYGLEWICZ, AND P. M. KORT (2014): “Willingness to wait under risk and ambiguity,” .
- EL-GAMAL, M. A. AND D. M. GREETHER (1995): “Are people Bayesian? Uncovering behavioral strategies,” *Journal of the American statistical Association*, 90, 1137–1145.
- FEHR, E. AND A. RANGEL (2011): “Neuroeconomic foundations of economic choice-recent advances,” *The Journal of Economic Perspectives*, 25, 3–30.
- FISCHBACHER, U. (2007): “z-Tree: Zurich toolbox for ready-made economic experiments,” *Experimental economics*, 10, 171–178.

- FREDERICK, S. (2005): “Cognitive reflection and decision making,” *The Journal of Economic Perspectives*, 19, 25–42.
- FRIEDMAN, D. AND J. RUST (1993): *The double auction market: institutions, theories, and evidence*, vol. 14, Westview Press.
- FUDENBERG, D., P. STRACK, AND T. STRZALECKI (2015): “Stochastic choice and optimal sequential sampling,” *Available at SSRN 2602927*.
- GABAIX, X., D. LAIBSON, G. MOLOCHE, AND S. WEINBERG (2006): “Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model,” *The American Economic Review*, 96, 1043–1068.
- GOLD, J. I. AND M. N. SHADLEN (2002): “Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward,” *Neuron*, 36, 299–308.
- GOLLIER, C. (2004): *The economics of risk and time*, MIT press.
- GREINER, B. (2015): “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, 1, 114–125.
- HOLT, C. A. AND S. K. LAURY (2002): “Risk aversion and incentive effects,” *American economic review*, 92, 1644–1655.
- JONES, M. AND R. SUGDEN (2001): “Positive confirmation bias in the acquisition of information,” *Theory and Decision*, 50, 59–99.
- KAHNEMAN, D. (2003): “Maps of bounded rationality: Psychology for behavioral economics,” *The American economic review*, 93, 1449–1475.
- KAHNEMAN, D. AND A. TVERSKY (1979): “Prospect theory: An analysis of decision under risk,” *Econometrica: Journal of the Econometric Society*, 263–291.
- KRAJBICH, I., B. OUD, AND E. FEHR (2014): “Benefits of neuroeconomic modeling: new policy interventions and predictors of preference,” *The American Economic Review*, 104, 501–506.

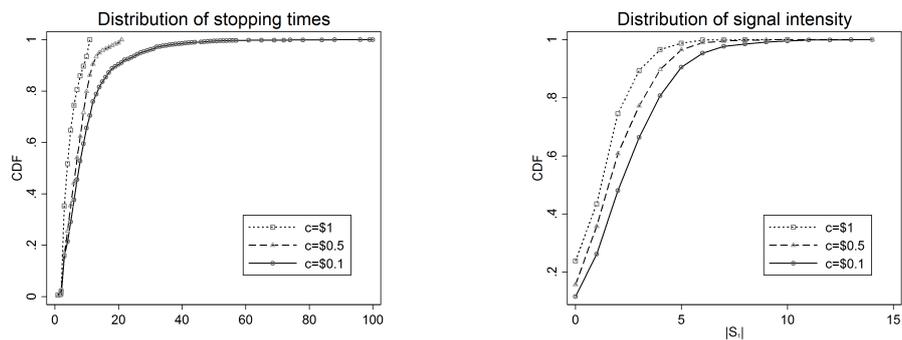
- KRUSE, T. AND P. STRACK (2015): “Optimal stopping with private information,” *Journal of Economic Theory*, 702–727.
- LOOMES, G. AND R. SUGDEN (1986): “Disappointment and Dynamic Consistency in Choice under Uncertainty,” *The Review of Economic Studies*, 53, 271–282.
- MARTIN, V., S. HURN, AND D. HARRIS (2012): *Econometric modelling with time series: specification, estimation and testing*, Cambridge University Press.
- MCDONALD, R. AND D. SIEGEL (1986): “The Value of Waiting to Invest,” *The Quarterly Journal of Economics*, 707–728.
- OPREA, R., D. FRIEDMAN, AND S. T. ANDERSON (2009): “Learning to wait: A laboratory investigation,” *The Review of Economic Studies*, 76, 1103–1124.
- RABIN, M. AND J. L. SCHRAG (1999): “First impressions matter: A model of confirmatory bias,” *Quarterly journal of Economics*, 37–82.
- RADNER, R. AND J. STIGLITZ (1984): “A Nonconcavity in the Value of Information,” *Bayesian models in economic theory*, 5, 33–52.
- RATCLIFF, R., P. L. SMITH, S. D. BROWN, AND G. MCKOON (2016): “Diffusion decision model: current issues and history,” *Trends in cognitive sciences*, 20, 260–281.
- TARTAKOVSKY, A., I. NIKIFOROV, AND M. BASSEVILLE (2014): *Sequential analysis: Hypothesis testing and changepoint detection*, CRC Press.
- TREICH, N. (1997): “Risk tolerance and value of information in the standard portfolio model,” *Economics Letters*, 55, 361 – 363.
- VIEFERS, P. AND P. STRACK (2014): “Too Proud to Stop: Regret in Dynamic Decisions,” *Available at SSRN 2465840*.

- WAKKER, P. P. (2010): *Prospect theory: For risk and ambiguity*, Cambridge University Press.
- WALD, A. (1945): “Sequential Tests of Statistical Hypotheses,” *The Annals of Mathematical Statistics*, 16, 117–186.
- WALD, A. AND J. WOLFOWITZ (1948): “Optimum character of the sequential probability ratio test,” *The Annals of Mathematical Statistics*, 326–339.
- WEBB, R. (2013): “Dynamic constraints on the distribution of stochastic choice: Drift Diffusion implies Random Utility,” *unpublished, New York University*.
- WEISS, C. H. (2008): “Thinning operations for modeling time series of counts: a survey,” *ASTA Advances in Statistical Analysis*, 92, 319–341.
- WILSON, A. (2014): “Bounded memory and biases in information processing,” *Econometrica*, 82, 2257–2294.
- WOODFORD, M. (2014): “Stochastic choice: An optimizing neuroeconomic model,” *The American Economic Review*, 104, 495–500.
- YARIV, L. (2002): “I’ll See It When I Believe It? A Simple Model of Cognitive Consistency,” .

# A Appendix

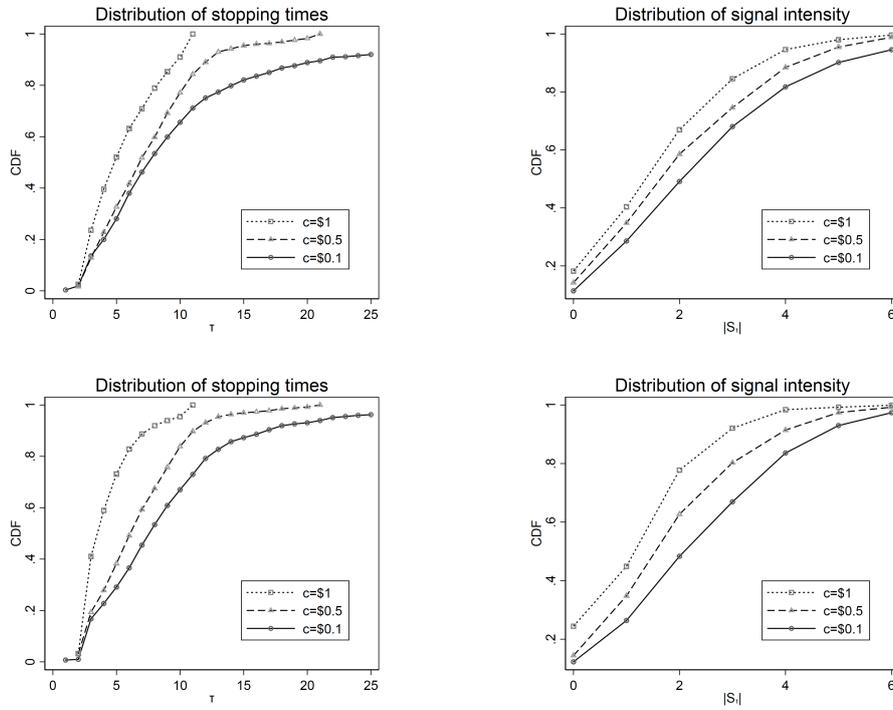
## A.1 Figures

**Figure 8:** Cumulative distribution of stopping times and signal intensity.



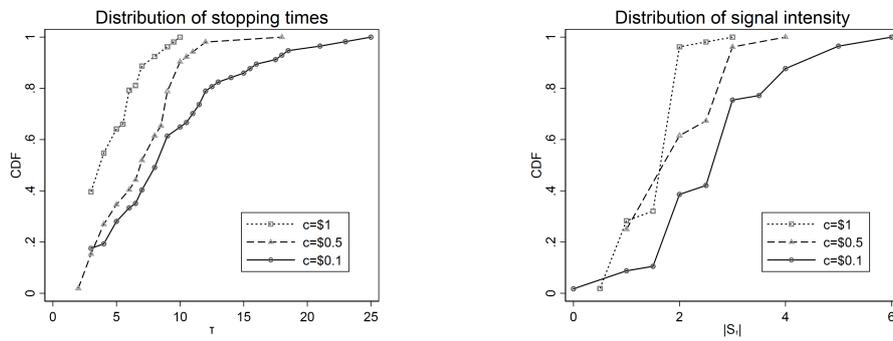
*Note:* This figure is similar to Figure 4 and shows cumulative distribution of stopping times and signal intensity without restriction on stopping times and signal intensity values.

**Figure 9:** Cumulative distribution of stopping times and signal intensity for the first 20 and last 20 periods.



*Note:* This figure is similar to Figure 4 and shows cumulative distribution of stopping times and signal intensity with restriction of stopping times below 26 and signal intensity below 6. The top part of the figure represents the cumulative distributions for the first 20 periods of the experiment; the bottom part represents them for the last 20 periods.

**Figure 10:** Cumulative distribution of median stopping times and signal intensity.



*Note:* This figure is similar to Figure 4 and shows cumulative distribution of median stopping times and median signal intensity.

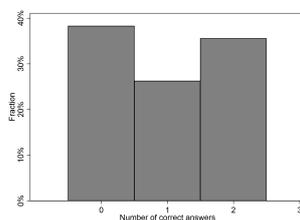
## A.2 CRT

This section summarises the results of the cognitive reflection test (CRT) that was run during the experiment. It is a test of cognitive ability, in the sense that it assesses the ability of a participant to control impulsive answers to simple questions. The three questions that were asked are borrowed from (Frederick, 2005) and are the following:

- 1 A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost? \_\_\_\_ cents
- 2 If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? \_\_\_\_ minutes
- 3 In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? \_\_\_\_days

We define the CRT score as the number of correct answers in the CRT. Figure 11 plots the distribution of the CRT scores. The average number of correct answers is 1.1. No subject managed to give the 3 correct answers. We cannot reject the hypothesis that the CRT has an impact on the rate of optimality in the task for both the strict and loose optimality conditions ( $p = 0.11$  and  $p = 0.83$  for an ANOVA test).

**Figure 11:** Plot of the distribution of the CRT score.



### A.3 Optimality rates with elicited risk parameters

In order to further investigate the impact of risk preferences on sampling decisions, we computed the optimality rate for each participants, according to their elicited risk preferences (using the Holt and Laury (2002) procedure). The results are displayed below.<sup>17</sup>

Overall, we find that using the individual risk aversion parameter has a very low explanatory power. In the \$0.10 treatment the strict and loose optimality rates are both below 20%. In the \$0.5 treatment, both the strict and loose optimality rates are below 35%. Finally, in the \$1 treatment, the loose optimality rate is relatively high, but remains significantly below 50% ( $p < 0.001$  for a WSR test)

**Table 5:** Summary of the estimation of optimality rate using the individual risk parameter.

	Optimality rate	
	strict	loose
\$0.10	14%	17%
\$0.50	18%	32%
\$1	24%	44%

### A.4 Screenshot

---

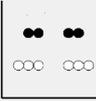
<sup>17</sup>Note that some observations are dropped from this estimation (around 18%) as the choices were inconsistent. For instance, some participants chose the last option in the procedure, where they chose \$1.5 with probability 1 compared to \$3 with probability 1

Figure 12: Sample Screen of the experiment

<b>Round :</b>	1
<b>if you find the right answer you can earn:</b>	20
<b>The price of a ball is:</b>	0.10
<b>You can still spend:</b>	9.20
<b>You can still earn up to:</b>	29.20

5  
 3

I guess the urn is A



URN A

?

I guess the urn is B



URN B



## B Experimental Instructions

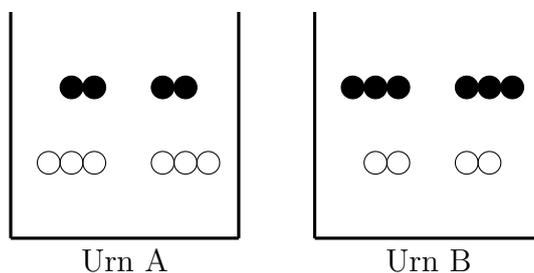
Welcome to our experiment! You will receive \$5 for showing up on time. Please read these instructions carefully and completely. Properly understanding the instructions will help you to make better decisions and, hence, to earn more money. If you read these instructions carefully and perform well in the experiment, you can earn more money (which will be paid out to you in cash at the end of the experiment).

Please keep in mind that you are not allowed to communicate with other participants during the experiment. If you do not obey this rule you will be asked to leave the laboratory and will not be paid. Whenever you have a question, please raise your hand; an experimenter will come to assist you.

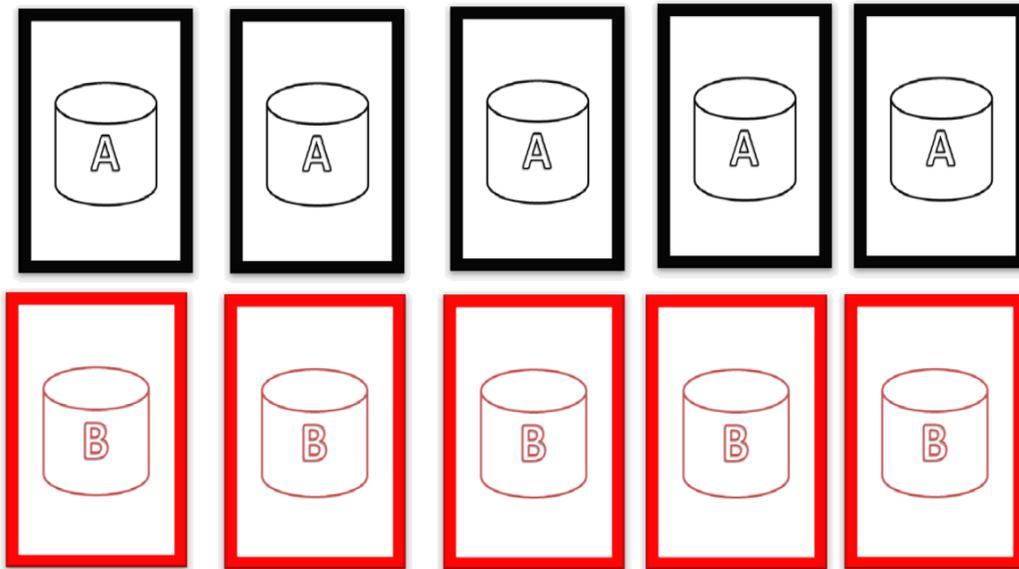
This experiment consists in multiple rounds. In each round you will earn money depending on your performance. At the end of the experiment one of the rounds will be randomly selected by the computer, and you will earn the payoff from this round.

### Setting

In this experiment there will be two urns: Urn A and Urn B. Urn A contains 4 black balls and 6 white balls. Urn B contains 6 black balls and 4 white balls. The following figure summarises the situation:



At the beginning of every round one of the two urns will be randomly selected by the computer, by picking up a card from the following deck:



## Your Objective

Your aim is to find what urn has been selected by the computer. In order to help you to reach your decision, balls will be automatically drawn from the urn. You are given a \$10 endowment every round, from which the cost of every ball will be withheld. The ball that has been drawn *goes back* into the urn after every draw. Each time a ball is drawn, you have to pay for it.

You can stop the process as soon as you think you know what urn has been selected.

All the relevant information will be displayed on the screen before any decision:

Round :	1
If you find the right answer you can earn:	20
The price of a ball is:	0.10
You can still spend:	9.20
You can still earn up to:	29.20

5  
 3

I guess the urn is A

URN A

?

I guess the urn is B

URN B

## Your Earnings

If you correctly guess what urn has been selected, you earn the payoff. If you are wrong, the payoff is zero. The cost of drawing a ball from the urn is withheld from your endowment of the corresponding round. Therefore, your earning for every round is as follows:

- If you are right: You *win* **\$20** and keep whatever you have not spent on drawing balls from your \$10 endowment.
- If you are wrong: You *win* **\$0** and keep whatever you have not spent on drawing balls from your \$10 endowment.

Before the end of the experiment, you will also enter a lottery, whose outcome will be added to your earnings. In this lottery, you can switch from A to B only once. If you want to change your mind you have to change your switching point.

Lottery A		Choice:	Lottery B	
10% chance of winning \$2	90% chance of winning \$1.60	A C F B	10% chance of winning \$3.85	90% chance of winning \$0.10
20% chance of winning \$2	80% chance of winning \$1.60	A C F B	20% chance of winning \$3.85	80% chance of winning \$0.10
30% chance of winning \$2	70% chance of winning \$1.60	A C F B	30% chance of winning \$3.85	70% chance of winning \$0.10
40% chance of winning \$2	60% chance of winning \$1.60	A C F B	40% chance of winning \$3.85	60% chance of winning \$0.10
50% chance of winning \$2	50% chance of winning \$1.60	A C F B	50% chance of winning \$3.85	50% chance of winning \$0.10
60% chance of winning \$2	40% chance of winning \$1.60	A C F B	60% chance of winning \$3.85	40% chance of winning \$0.10
70% chance of winning \$2	30% chance of winning \$1.60	A C F B	70% chance of winning \$3.85	30% chance of winning \$0.10
80% chance of winning \$2	20% chance of winning \$1.60	A C F B	80% chance of winning \$3.85	20% chance of winning \$0.10
90% chance of winning \$2	10% chance of winning \$1.60	A C F B	90% chance of winning \$3.85	10% chance of winning \$0.10
100% chance of winning \$2	0% chance of winning \$1.60	A C F B	100% chance of winning \$3.85	0% chance of winning \$0.10

At the end of the experiment one of the rounds is selected, and you will earn the payoff from this round:

**Results of the experiment and Payment**

<u>Decisions</u>	
Number of Right Decisions	1
The computer selected round:	1
In round 1 you were:	wrong
<u>Payment in AUD:</u>	
Show up fee:	\$5.00
Payoff from the selected round:	\$6.70
Payoff from the selected lottery:	\$1.60
Final profit:	\$13.30