Paying is Believing: The Effect of Costly Information on Bayesian Updating

Pedro Robalo
Max Planck Institute for Research on Collective Goods

Rei Sayag
Universitat Pompeu Fabra*

June 2015

Abstract

The difficulties posed by Bayesian updating are recognized across many domains. In this paper we explore whether individual belief updating is affected by the cost of information. Our conjecture is that this effect should be observed if individuals are prone to the sunk cost fallacy. We design an experimental environment where subjects perform a belief updating task after receiving useful and identical information on the state of the world. Our treatments vary the way in which information is made available to subjects. We find a systematic effect of the cost of information on belief updating. Subjects overweigh costly information relative to free information, which results in a 'push' of beliefs towards the extremes. The cost-driven shift can lead to posterior beliefs more attuned with Bayesian updating. We argue that an intensification of the representativeness bias is the most likely explanation of our results.

Keywords: information; Bayesian updating; decision making and risk; sunk cost; heuristics and biases.

JEL codes: C91; D81; D83.

*Sayag (corresponding author): Universitat Pompeu Fabra, Ramon Trias Fargas, 25-27, Barcelona, Spain. E-mail: sayag@ese.eur.nl. Robalo: Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53113, Bonn, Germany. E-mail: robalo@coll.mpg.de.

1 Introduction

The process of information acquisition plays a crucial role in supporting decisions made by individuals and organizations. As new information is collected, better understanding of the environment is achieved and improved decisions can be made. Nonetheless, it is well known that the integration of new information with prior knowledge can result in biased beliefs (Alberoni, 1962; Kahneman and Tversky, 1972; Hammerton, 1973) and bad decisions (Charness et al., 2007; Redlawsk, 2002; Camacho et al., 2011). This paper investigates how the cost of new information affects its use in the decision making process. For example, should it matter for an executive's decision whether she received market analysis cheaply from in-house analysts or from a more expensive outside consultancy firm? Similarly, should it matter for a consumer's purchasing decision whether his knowledge of a product's quality is the result of a lengthy search process or based on a quick judgment? Both examples contain an element of sunk cost: once information is received, conditional on quality, the cost should have no effect on the resulting beliefs and decisions. In more general terms, the behavior of a rational decision maker should not depend on the cost of information, all else equal. However, we conjecture otherwise: decision makers might put a higher weight on costly information, regardless of its quality, affecting their decisions.

Underlying our conjecture is the possibility that decision makers fall prey to the sunk cost effect (Thaler 1980), and "use" information relatively more when it comes at a cost.[1] If this is true, a possible implication is that the cost of information can be used to steer decision makers' behavior in the preferred direction. We contribute to the accumulated evidence on the sunk cost fallacy by investigating the existence of sunk costs in a scenario of decision making under risk. If a relationship between information costs and decision making exists, a follow-up question is whether it leads to better decisions.

We set out to investigate these matters using a laboratory experiment. Field data is likely to be contaminated by serious selection issues: individuals who choose to acquire information in the field are likely to differ along several dimensions from individuals who choose not to do so. The laboratory allows us to correct for these selection issues in two ways. First, we are able to disentangle selection from sunk cost effects by imposing the cost of information on subjects. This is something that is easily done in the laboratory, but arguably difficult to implement in the field.[2]

[1] According to Thaler (1980): "paying for the right to use a good or service will increase the rate at which the good will be utilized, ceteris paribus. This hypothesis will be referred to as the sunk cost effect."

[2] There exist field tests of sunk cost effects in product use (Arkes and Blumer 1985, Ashraf et al. 2010 and Cohen and Dupas 2010, for example), but doing so with respect to information is arguably more complicated. In particular,

Second, the laboratory allows us to assess the extent to which individuals value information and are able to use it. To the best of our knowledge, this paper provides the first experimental evidence on the manner in which the cost of information affects belief updating behavior.

The sunk cost fallacy's main prescription is that only marginal costs and benefits should matter for decision-making. The vintage normative prescriptions (e.g. "don't push yourself through a movie that you are not enjoying") are among the first lessons that business and economics students are exposed to. However, the sunk cost fallacy still seems to plague many courses of action, be it continuing a failed relationship because one has already invested many years in it or a failure to withdraw from a lost war because of an extensive death toll. Thaler (1980) puts forward a compelling loss aversion-based rationale for why people fall prey to the sunk cost fallacy. Given the convexity of the utility of losses, a decision-maker facing a risky investment has an incentive to recover an incurred loss because the increase in utility of a gain will be larger than what a further comparable loss would entail. Eyster (2002) presents a taste-for-consistency-based explanation of the sunk cost fallacy. In face of sequential decisions under risk, decision makers trade off revenue-maximizing choices for consistency-maximizing ones, i.e. a decision maker gives up revenue today if this choice makes yesterday's decision look optimal.

Despite the abundance of casual and anecdotal evidence, the literature's verdict on the sunk cost fallacy is still mixed. The pioneering field experiment of Arkes and Blumer (1985) found that granting a random discount for a theater season ticket significantly decreases attendance, which is evidence of a sunk cost effect. Drawing inspiration from this study, Ashraf et al. (2010) and Cohen and Dupas (2010) test for selection and sunk cost effects in the pricing of health products in the developing world; they find weak evidence of sunk cost effects. Other tests with field data have also produced mixed evidence: Staw and Hoang (1995) find considerable sunk cost effects in the drafting of National Basketball Association players (a result later corroborated by Camerer and Weber 1999), while Borland et al. (2011) find no such effects for the Australian Football League. In an education setting, Ketel et al. (2015) find that university students are not prone to sunk cost effects, as granting a random tuition fee discount does not influence student participation or performance. On the other hand, Herrmann et al. (2015) show that delegating decisions to a computer in bidding-fee auctions is associated with a reduction of sunk cost effects due to a lower behavioral investment from bidders. Ho et al. (2015) find that an exogenous variation in car prices

measuring product usage (a theater season ticket, a bottle of water disinfectant and bed nets, respectively for the cited works) is easier than measuring information usage.

leads to different usage rates by Singaporean drivers, which hints at a sunk cost effect.

The experimental laboratory evidence on the sunk cost fallacy is also mixed. On the one hand, using a search environment specifically designed to observe sunk cost effects, Friedman et al. (2007) find that experimental subjects are surprisingly consistent with optimal behavior, falling prey to the sunk cost fallacy only occasionally. On the other hand, in an industrial organization setting both Offerman and Potters (2006) and Buchheit and Feltovich (2011) find that sunk costs influence pricing decisions. Cunha and Caldieraro (2009) show that sunk costs not only affect decisions over material investments, but also behavioral ones, i.e. those which stem from the cognitive effort invested in a task. They show that subjects are more likely to switch to a slightly better alternative if the sunk level of effort was low. However, an attempt at replicating these findings was not successful (Otto 2010).

The work of Gino (2008) is methodologically close to this paper, but focuses on the role of the cost of advice. Subjects in her experiment answer trivia questions, for which they can use free or paid advice. Subjects who pay for advice incorporate it significantly more often in their decisions than those who obtain advice for free. From the competing explanations, the author shows that sunk cost effects drive the results. However, the application of her results to standard models of decision making under risk is limited by the the fact that subjects' prior and posterior beliefs are not known to the experimenter. Our experimental design allows for these measurements and therefore uncovers the conditions under which costly information influences belief updating. Furthermore, we are able to investigate the channel through which sunk costs operate and test for potential selection effects, both within- and between-subjects.

Our study investigates the impact of the cost of information in a setting where subjects have to make a decision under risk. Information is provided in a way that can help them reduce uncertainty in a Bayesian fashion, and therefore our work relates to a long literature in economics and psychology that deals with optimal decision making under risk, as well as the associated heuristics and biases (see DellaVigna 2009 for a review). In particular, we are interested in knowing whether the cost of information can play a role in dampening some of the traditional biases or interact with some common heuristics. To be sure, the verdict on whether "man is a Bayesian" is still out. When combining information on prior probabilities of possible states of the world with informative state-dependent signals, three main inter-related phenomena have been observed (see Camerer 1995 for a detailed overview). First, individuals often exhibit conservatism in their choices, failing to use new information to the extent normatively prescribed by Bayes' formula (e.g. Eger and Dickhaut

1982). Second, there is a systematic tendency for individuals to neglect prior probabilities in their judgment, often referred to as "base rate neglect" (Koehler 1996). Third, when the signal is representative of one of the states, the tendency to overweigh the signal's information content is exacerbated. This heuristic is known as "representativeness".[3] For example, if a decision maker draws a sample which *exactly* matches the distribution of a certain population, she will tend to overweigh the probability that this sample comes for that population (in which case the heuristic is referred to as "exact representativeness").

Early evidence (e.g. Kahneman and Tversky, 1972 and 1973) showed that representativeness was a serious and systematic bias. A number of experiments by David Grether (1980, 1992; El-Gamal and Grether 1995) produced more optimistic evidence: subjects do use representativeness (especially when it is "exact"), but behavior is not always far from Bayesian. Even though experimental subjects prove not to be perfect Bayesians, El-Gamal and Grether (1995) showed that a Bayesian decision rule receives the strongest empirical support, followed by representativeness. Experimental market tests of this heuristic (Duh and Sunder 1986 and Camerer 1987) have shown that behavior converges to Bayesian over time and that the observed deviation is mostly explained by representativeness. In sum, with respect to conservatism, base rate neglect and representativeness, the accumulated evidence seems to show that "base rates are underweighted in some settings but sample information is underweighted in others. Base rates are incorporated when they are salient or interpreted causally." (Camerer 1995).

Building upon these conclusions, we ask a natural question: can the cost of information influence the extent to which conservatism, base rate neglect and representativeness prey on decision makers? If that is the case, the cost of information can be used to dampen some of the shortcomings associated with decision making under risk. We seek to establish an existence result which would allow for further context-specific investigations where information cost is the control variable.

In our design each subject has to make a number of decisions with state-dependent payoff consequences. There are two states of the world with known and constant priors. Subjects sometimes have the opportunity of receiving additional information by drawing a sample (a "ball") from a state-dependent lottery (an "urn"). Our treatments change the way in which this information is made available: in *Free* it has no cost, while in *Costly* it is accessible if purchased. A treatment where the cost is imposed on subjects (*Compulsory*) corrects for selection while leaving the role

[3] According to Kahneman and Tversky (1972): "this heuristic evaluates the probability of an uncertain event, or a sample, by the degree to which it is: (i) similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated."

of cost intact. Moreover, and for all treatments, subjects subsequently complete a reduced version of the three treatments. Observing subjects' revealed demand for information allows us to further analyze the role of individual heterogeneity on selection and updating behavior.

Our results show that there are systematic biases in individual belief updating. Paying for information in the *Costly* and *Compulsory* treatments leads to an over-weighting of newly obtained information, which leads to moves in the posterior that are more extreme than in the *Free* treatment. This pattern is explained by a sunk cost effect, as the only difference between the two former treatments and the latter is the cost charged for information. These results cannot be explained by selection, as we detect no significant differences between *Compulsory* and *Costly*. Moreover, individual heterogeneity does not explain the overall pattern, which reinforces the sunk cost explanation. Regarding decision optimality, choices with costly information can lead to better or worse outcomes, depending on whether subjects fall short or exceed the Bayesian updating benchmark when information is provided for free. The use of costly information is beneficial if belief updating after free information corresponds to a situation of Bayesian under-updating. Costly information leads subjects to put a higher weight on newly obtained information.

2 Experimental Design

Each subject is presented with a 'book bags and poker chips' belief updating task (see Table 1). In this task one of two states of the world can occur (**L**eft or **R**ight), with known probabilities: $p \equiv \Pr(L)$ and $1 - p \equiv \Pr(R)$. The subject is presented with a state-dependent lottery, which provides an informative signal. This lottery is framed as an urn containing five balls, some black and some white, depending on the state of the world. The belief updating task is repeated 40 times, with constant prior probabilities and urn composition.[4]

	State of the world	
	Left	Right
Prior probability	0.4	0.6
Urn composition	●○○○○	●●●○○

Table 1: Priors and signal distribution in the two states of the world.

Individual posterior probability beliefs are elicited using the following incentivized mechanism.

[4]We chose not to implement symmetric priors since the task could become trivial (Camerer 1987) or invite the use of "obvious" (but possibly wrong) heuristics.

We implement a scoring function that consists of state-dependent payoff correspondences, depicted in Figure 1.[5] Subjects choose a number between 0 and 100 in steps of 0.5. If the state is L (R) the optimal decision is 20 (80). A posterior probability belief can be recovered directly from a subject's decision. For example, choosing 60 is optimal if a subject believes that $p = 0.38$. The information on the two state-dependent payoffs is made available to subjects in three distinct ways: on screen (subjects can interactively learn the state-dependent payoffs that result from any particular decision at all times using a slider bar), in graphical format and in table format (both provided in the paper instructions). We believe that this choice framework is simpler and more intuitive for experimental subjects than other commonly used probability elicitation mechanisms (e.g. Becker-De Groot-Marschak mechanisms). [6]

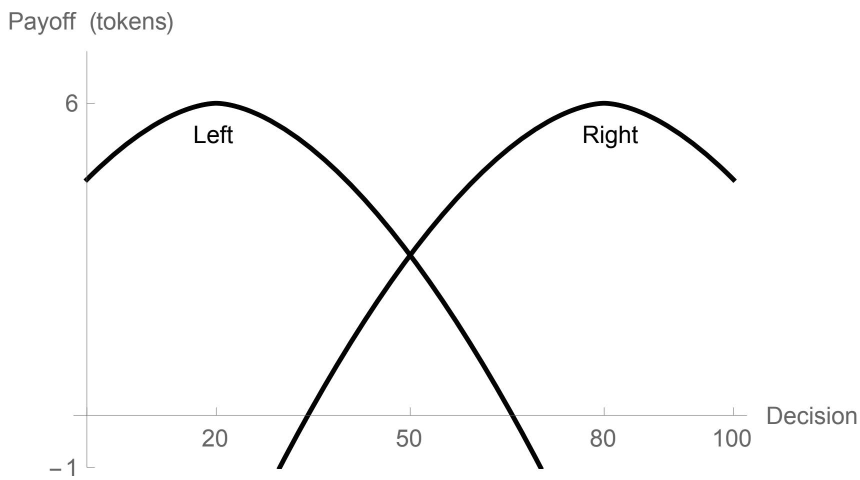

Figure 1: Scoring function. The left (right) component applies in the corresponding state.

2.1 Treatments

We implement our treatments by varying the way in which information is made available to subjects. More specifically, we vary the cost of a ball draw and whether payment is mandatory or optional.

[5]This state-dependent payoff function is a proper scoring rule, i.e. it induces truth-telling for risk neutral subjects. The fundamental reason that led us to implement this scoring rule and not one of the commonly used proper scoring rules, like the quadratic scoring rule (reviewed by e.g. Schlag et al. 2014), is the fact that many of these do not provide a substantial incentive to update beliefs unless radical moves in the posterior are observed. In other words, we opted for a scoring rule that is steep enough in the region where probability updating takes place. A common problem with proper scoring rules is that risk attitudes may play a role in the observed choices. However, this is not problematic in our setting as risk attitudes influence subjects' decisions identically across treatments. Nonetheless, we statistically control for risk attitudes in our analysis.

[6]See Appendix A for a detailed description of the choice environment and derivation of optimal decisions according to the normative model. See the "Experiment Emulation" files in the supplemental material for an emulation of the belief elicitation task.

In the *Free* treatment the ball can be drawn from the urn at no cost. In *Costly* a ball can be drawn at a cost. In *Compulsory* the price of drawing a ball is imposed on subjects (subjects are told "a ball has been drawn for you"). In *Costly* a subject buying a ball observes it automatically while in *Free* and *Compulsory* subjects can choose whether to see the drawn ball or not. In all treatments, there is a 50% chance that subjects can draw a ball from the urn in each period of *Free* and *Costly*, or that a ball is drawn for them in *Compulsory*. If information were available in all periods we would run the risk of subjects automatically discounting the costs of information to be incurred at the beginning of the experiment, which would dissolve the psychological impact of the imposed cost. This also forces subjects to experience decisions without information, which provides us with individual decisions without information - a likely anchor for decisions when information is available. In *Costly* and *Compulsory* the information is priced at roughly 60% of the expected gain if expected utility maximization with Bayesian updating is performed by a risk- and loss-neutral decision maker. Note that the quality of the information is the same regardless of cost.

After subjects complete the 40-period decision task, they are presented with an extra decision task. This task is also incentivized, does not vary across treatments and is identical to the first one in all respects except for a different parametrization and the fact that it consisted of 3 sequences of 10 periods. The data from the extra decision task was used in further controlling for selection effects. This analysis is carried out in detail in Appendix B.

In order to control for risk attitudes and demographic characteristics in the statistical treatment of the data, we end the experiment with the Charness-Gneezy-Potters task for risk attitude elicitation (Gneezy and Potters 1997, Charness and Gneezy 2010) and a questionnaire.[7]

Paper instructions were distributed at the beginning of the experiment, which subjects were asked to read silently.[8] The experiment started after all subjects had finished reading the instructions. A set of practice questions to test the understanding of the experiment was then administered. A similar set of instructions was distributed before the extra decision task. In the experiment all values are expressed in tokens, which are converted at an exchange rate of 0.75 Euro per token. Subjects were paid for six randomly determined periods.

[7]The risk attitude elicitation task consists in asking subjects how they wish to allocate an endowment of three tokens between a safe account and an account that multiplies the invested amount by a factor of 2.5 with 50% probability, and destroys the money with 50% probability. The questionnaire asked whether subjects had math in high school, how many math courses they had completed thus far at the university level, as well as their gender, age, and study major.

[8]A transcript of the instructions can be found in Appendix C.

3 Experimental Results

The experimental sessions were run at the CREED laboratory of the University of Amsterdam between February and May 2012. A total of 166 subjects participated in 8 sessions, recruited online from a subject pool of students at the University of Amsterdam. No subject participated in more than one session. Fifty-five per cent of the participants were male and 57% were Business or Economics majors. The typical session lasted 1 hour and 20 minutes and the average earnings were 24 Euro (which includes a show-up fee of 7 Euro). The experiment was programmed with z-Tree (Fischbacher 2007). Subsection 3.1 analyzes the difference in belief updating across treatments, while subsection 3.2 investigates the possible channels through which sunk costs can affect belief updating in our setting.

Table 2 provides a summary of descriptive statistics for the collected data. Differences in individual traits are not statistically significant across treatments (Pearson's chi-square test $p > 0.17$). Average period payoff refers to the gross payoff, i.e. not including information costs. There is no significant differences in the average payoff between treatments. The percentage of information seen refers to the fraction of times subjects chose to observe information when it was available. Naturally, when information was costly and optional fewer subjects chose to observe it (two-sided Mann-Whitney-Wilcoxon rank-sum test, MWW hereafter: p=0.00 and p=0.04, respectively). The *Costly* treatment thus shows significantly fewer information views than the *Free* and *Compulsory* treatments. We observe that subjects chose not to see information (draw a ball) sometimes, even when it was free or already paid for. This is to a great extent the consequence of subjects experimenting with drawing and not drawing a ball.

	Free	*Costly*	*Compulsory*
N	65	65	36
Risk	2.06	1.94	2.10
Math courses	2.38	2.95	2.56
% Female	37%	48%	56%
Average period payoff	3.10	3.08	3.10
% Information seen	79%	55%	74%

Table 2: Summary statistics

3.1 Treatment effects

We begin the analysis with a description of mean treatment posteriors. The Bayesian posterior benchmarks are 0.18, 0.57 and 0.40 for a black ball ("Black"), a white ball ("White") and no information ("No Draw"), respectively. The corresponding mean individual posteriors across all treatments are 0.23, 0.64 and 0.42. Table 3 presents mean posterior beliefs by information condition and treatment. Focusing on treatment differences within information conditions, we observe that, for both White and Black, there exist significant differences between the *Costly* and *Free* treatments (MWW: $p = 0.01$ and $p = 0.03$, respectively), and between the *Compulsory* and *Free* treatments (MWW: $p = 0.00$ and $p = 0.02$, respectively). No significant differences are found between the *Costly* and *Compulsory* treatments in any information condition. Without a draw from the urn there are no significant differences between the treatments. The shift in posterior beliefs between both the *Costly* and *Compulsory* treatments and the *Free* treatment is thus significant only after a ball draw. It is a downward shift after a black draw and an upward one following a white draw.

	Black	White	No Draw
Free	0.27	0.60	0.43
	(0.02)	(0.02)	(0.01)
Costly	0.21	0.65	0.42
	(0.02)	(0.02)	(0.01)
Compulsory	0.19	0.70	0.42
	(0.03)	(0.03)	(0.02)
Benchmark	0.18	0.57	0.40

Table 3: Mean decisions by treatment and information condition. Note: Standard errors in parentheses.

Figure 2 presents the cumulative distribution functions of individual posteriors by information condition. We define the individual posterior $p_{i\phi t}$ as the elicited posterior probability of subject i in information condition $\phi \in \{Black, White, No\ Draw\}$ in period t. The mean posterior $\bar{p}_{i\phi}$ of subject i in information condition ϕ is:

$$\bar{p}_{i\phi} = \frac{\sum_{t=1}^{40} p_{i\phi t}}{n_{i\phi}} \tag{1}$$

where $n_{i\phi}$ is the number of periods (out of 40) in which information condition ϕ materialized.

As in Table 3, a shift in posteriors between the *Free* treatment and the *Costly* and *Compulsory* treatments is observable after a ball draw. Posterior distributions in *Costly* and *Compulsory* are first-order stochastically dominated by the posteriors in *Free* after a black draw, and are first-order stochastically dominant after a white draw. The differences in distributions are all significant at the 10% level (two-sample Kolmogorov-Smirnov test, see Table 4).

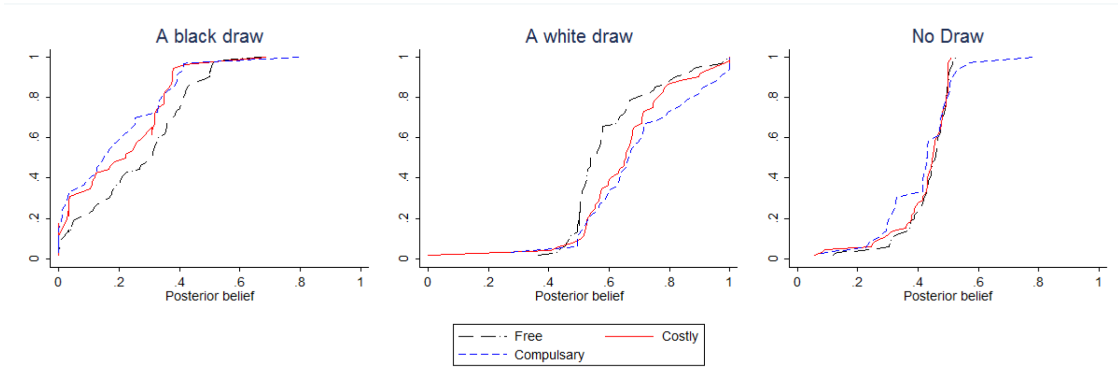

Figure 2: CDFs of individual posterior belief

	Black	White	No Draw
Free vs. *Costly*	0.05	0.01	0.95
Free vs. *Compulsory*	0.07	0.00	0.2
Costly vs. *Compulsory*	0.76	0.55	0.34

Table 4: p-values for Kolmogorov-Smirnov tests of posterior belief distributions

From Figure 2 and Tables 3 and 4 it is clear that the availability of costly information "pushes" subjects' posterior probability beliefs to the extremes. Subjects who incur a cost tend to hold more extreme posteriors vis-a-vis those who did not. As a check on these results, we also estimate a linear model using the mean posterior $\bar{p}_{i\phi}$ as the dependent variable and treatment dummies, gender, math courses and risk attitude as independent variables. We find significant and systematic effects of the *Costly* and *Compulsory* treatments on posterior beliefs.[9]

The overall evidence shows that subjects who pay for information behave differently than those who do not, which supports the existence of a sunk cost effect in belief updating. Thaler's (1980) description of the mechanism behind the sunk cost phenomenon provides an explanation for the observed data. A subject paying for information, found in a loss state, has a higher marginal payoff benefit than a subject receiving information at no cost. If effort is costly, a paying subject should exert higher effort than a non-paying subject. Higher effort may change the way in which information is used; in our setting this change translates into a stronger weighting of new information in belief updating. An equivalent manner in which sunk costs can lead to this change is by increasing the relative salience of a ball draw to a subject who just paid for it. Following what was hinted at

[9]See Table 9 in Appendix B.

by Koehler (1996), the change in the salience of information (in this case due to the cost paid for it) may affect the use of information. We establish our first result:

Result 1: *Paying for information affects individual posterior probability estimates in a systematic manner. Individuals who incur a cost overweigh the newly acquired information relative to individuals who do not.*

A natural concern is that selection is driving the difference between the *Costly* and *Free* treatments. That is, the two samples are not identical as subjects who choose to pay for information may differ in many dimensions from those who are only willing to observe free information. The mean posterior in the *Free* treatment can then be perceived as a weighted average of two sub-groups (those who are willing to pay for information and those who are not), while the mean posterior in the *Costly* treatment after a ball draw is the outcome of only the former group's decisions. However, selection can not play a part in the differences between the *Free* and the *Compulsory* treatments, as evidenced by the similar information acquisition rates in Table 2. Any difference in posteriors should thus be attributed to the difference in cost between these two treatments. In addition to this identification mechanism, Appendix B elaborates on the absence of selection effects using data from the extra decision task.

We now turn to decision optimality across treatments, where we define optimality as the absolute distance from the Bayesian posterior benchmark. Table 3 presents the treatment means by information condition. We observe that the mean posteriors in the *Free* treatment is the most optimal after a white draw, but the least optimal after a black draw, for which the most optimal outcomes are obtained in *Compulsory*.

Result 2: *Costly information can lead to a more or less optimal use of information in a belief updating task.*

We conclude that, depending on the state of the world, costly information may improve or worsen a subject's performance. This is a result of the "push" towards the extremes that costly information induces. After a black draw, since subjects tend to under-update new information in the *Free* treatment, costly information leads to a more optimal decision. In case of a white draw, where subjects over-update, costly information leads to a less optimal decision.[10]

[10] Considering Result 2 in light of Thaler's (1980) explanation of sunk cost, it might be surprising at first look that increased effort can lead to a less optimal result. Still, similar observations in the literature exist which demonstrate that more effort can lead to lesser performance (e.g. Camerer and Hogarth 1999, Ariely et al. 2009, Leuven et al. 2011).

3.2 Interpreting the effect of cost

Our experimental design is devised to detect the existence of sunk cost effects on the use of information, which we have shown to exist. In this subsection we aim to explore the underlying channels through which the cost of information affects belief updating. We seek to connect our findings to the biases in belief updating discussed in Section 1, namely representativeness, conservatism and base-rate neglect. Note that our experiment does not single out an interpretation but allows us to point at a likely candidate.

Representativeness refers to an individual's tendency to overweigh new information, biasing her posterior belief away from the prior. When observing a white ball draw, this bias leads an individual to overweigh the likelihood that this ball draw signals the state of the world is Left. A similar overweighting of the Right state of the world occurs if the ball is black. Representativeness thus results in posteriors which are further away from the prior than the Bayesian posterior benchmark dictates. Conservatism has the opposite effect on posterior belief formation. It refers to an individual's tendency to underweight new information, thus biasing her posterior belief towards the prior. Base rate neglect pertains to the tendency of an individual to underweigh the prior when receiving new information, i.e. with two possible states of the world, the individual has a prior belief that is closer to 0.5 than it actually is. Note that this is *not* the same as representativeness. Since the prior in our design indicates that Left is less likely than Right, this bias brings subjects to over-update the probability of Left after any ball draw.

An increase in the strength of base rate neglect due to costly information could be a natural channel for the effect of cost on posteriors. A subject who focuses more on the new information she just paid for automatically discounts the underlying prior. This explanation fits the observed behavior after a white draw but not after a black draw. If base rate neglect increased with a costly ball draw, the average posterior after a costly black draw should be higher than after a free black draw, while we observe the opposite.

The pattern observed in our data seems to fit with the representativeness heuristic rationale. If paying for information engenders or intensifies representativeness, a subject observing a white ball perceives it to be more representative of Left (out of 5 balls, 4 are white in Left and 2 are white in Right), and a black ball to be more representative of Right. This would directly lead to more extreme decisions.

We investigate the representativeness explanation using a parametric estimation of a simple model of information updating that builds upon a Bayesian updating framework used by Grether (1980, 1992) to assess representativeness and conservatism, and by Gonzales and Wu (1999) and Holt and Smith (2009) to evaluate probability weighting. The individual belief that the state of the world is Left after Black can be written as:

$$\Pr(Left|Black) = \frac{(\Pr(Black|Left))^{\eta} \cdot (\Pr(Left))^{\gamma}}{(\Pr(Black|Left))^{\eta} \cdot (\Pr(Left))^{\gamma} + (\Pr(Black|Right))^{\eta} \cdot (\Pr(Right))^{\gamma}} \tag{2}$$

If both $\gamma = 1$ and $\eta = 1$ the individual is Bayesian. The parameter $\gamma \in [0, 1]$ represents the strength of base rate neglect. If $\gamma = 1$ there is no base rate neglect, while as γ goes to zero the strength of base rate neglect increases, i.e. the prior probabilities become more equal. Similarly to Camerer (1987), if $\gamma = 0$ the individual perceives the prior probabilities of the two states of the world as identical. The representativeness/conservatism bias is brought about by the parameter $\eta \in [0, \infty]$. If $\eta > 1$ the individual overweighs new information such that representativeness bias exists. If $\eta < 1$ the opposite occurs and conservatism bias is observed. We can write the individual updated odds ratio after a black ball draw as:

$$\frac{\Pr(Left|Black)}{\Pr(Right|Black)} = \left(\frac{\Pr(Black|Left)}{\Pr(Black|Right)}\right)^{\eta} \cdot \left(\frac{\Pr(Left)}{\Pr(Right)}\right)^{\gamma} \tag{3}$$

A similar expression can be written for a white draw. To estimate this model while allowing for treatment effects we use a log-linear least squares specification:[11]

$$\ln\left(\frac{\bar{p}_{i\phi}}{1 - \bar{p}_{i\phi}}\right) = \alpha + (\beta_1 + \beta_2 \cdot D_{Costly} + \beta_3 \cdot D_{Compulsory}) \cdot \ln\left(\frac{\Pr(Draw|L)}{\Pr(Draw|R)}\right) \tag{4}$$

where $\bar{p}_{i\phi}$ is the posterior probability of Left for information condition ϕ and subject i, as given in equation (1). D_{Costly} and $D_{Compulsory}$ are treatment dummies, while $\eta = \beta_1 + \beta_2 \cdot D_{Costly} + \beta_3 \cdot D_{Compulsory}$ and $\alpha = \gamma \ln(\Pr(Left) / \Pr(Right))$. The parameter η is decomposed into three components in order to capture treatment effects via the likelihood ratios in the different treatments. The prior odds are included in the constant term since we cannot estimate γ due to the lack of variation in the prior odds of the state of the world.

Column (1) in Table 5 presents the estimation results of equation (4). We cannot reject that

[11] For posteriors with no ball draw equal likelihood are assumed $\Pr(Black|Left) = \Pr(Black|Right) = 0.5$. Following Grether (1992) and Holt and Smith (1999), we change inferred probabilities of 0 and 1 to 0.01 and 0.99, respectively.

	Model	
	(1)	(2)
β_1 (Likelihood ratio)	1.07^{***}	1.08^{***}
	(0.13)	(0.12)
β_2 (Likelihood ratio x *Costly*)	0.44^{**}	0.43^{**}
	(0.21)	(0.20)
β_3 (Likelihood ratio x *Compulsory*)	0.74^{***}	0.75^{***}
	(0.25)	(0.28)
Costly	–	-0.07
	–	(0.11)
Compulsory	–	0.04
	–	(0.14)
Constant	-0.25^{***}	-0.24^{***}
N	460	460

Table 5: Representativeness bias and treatment effects. Note: Each observation corresponds to the average individual decision in a specific information condition. The number of observation is lower than $166\times3=498$ since for some subjects there are no observations for all information conditions. Clustered standard errors used. $^{***}/^{**}/^{*}$ indicates significance at the 1%/5%/10% level.

β_1 differs from 1 (Wald, $p = 0.53$), suggesting that there is no representativeness or conservatism bias in the *Free* treatment, i.e. $\eta = 1$. The interaction coefficients β_2 and β_3 are both positive and significantly different than zero. This suggests that the effect of costly information is to increase representativeness bias. Column (2) presents an estimation with independent treatment dummy variables. This allows for the possibility that the *Costly* and *Compulsory* treatments have an effect on individual decisions through channels other than representativeness (i.e. through the scope of base rate neglect γ). We do not find a significant effect of the treatment dummies on inferred probability odds. Our estimation results support the preceding qualitative discussion on the possible channels through which sunk costs affect individual belief updating behavior.

Result 3: *Costly information intensifies the level of representativeness bias.*

While average treatment effects are this paper's main focus, we observe a large heterogeneity in individual posterior beliefs, as depicted in Figure 2. We complement the analysis above by examining how our treatments affect individual biases in belief updating. Figure 3 presents individual estimations of η, the parameter capturing representativeness and conservatism, by treatment. The parameter η is obtained by estimating the model given by equation 4 for each individual over the 40 periods of the decision task.[12] We find there is consistency in individual updating behavior, or that individual subjects tend to over- or under-update their beliefs across information conditions.

[12]Subjects who do not observe a black ball and a white ball at least once (18% of subjects) are discarded.

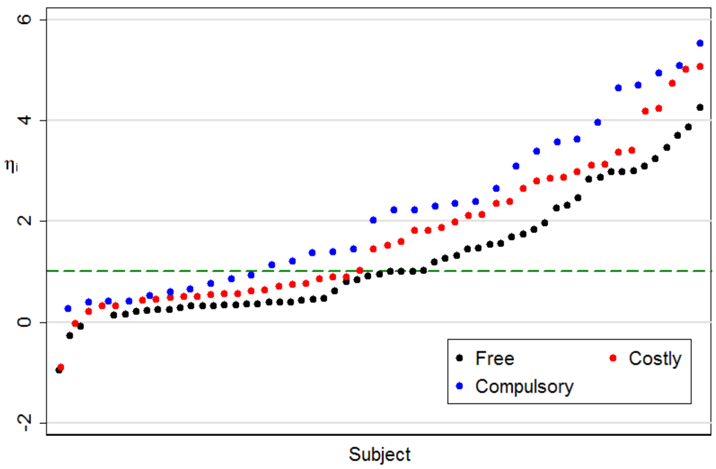

Figure 3: Estimated likelihood probability weight (η) for each individual, by treatment. The dashed green line is drawn at $\eta = 1$.

Across all treatments, 55% of subjects exhibit representativeness and the rest conservatism. It should also be noted that 47% of individual coefficients μ are not significantly different from 1. The large variation in departure from the Bayesian benchmark shown in Figure 3 is quantitatively similar to the result found in Holt and Smith (2009). The fractions of subjects with a representative bias is higher with costly information than without. These fractions are 0.49, 0.55, 0.67 in the *Free*, *Costly* and *Compulsory* treatments, respectively. The differences in distribution of η between the *Free* treatment and the *Costly* and *Compulsory* treatments are significant at the 10% level (two-sample Kolmogorov-Smirnov test).

Taken together, our results show that paying for information leads subjects to assign a higher weight to it. This pattern can be explained by a sunk cost argument: costly information is "used" more than free information. With respect to the shortcomings associated with Bayesian updating, our sunk cost effect manifests itself as an intensification of representativeness, i.e. the tendency of subjects to overestimate the likelihood of a state of the world when the information they receive matches more closely what is typical in that event. In other words, a subject who purchases costly information believes it more than if this information had been free.

4 Conclusion

This paper sets out to explore how a decision maker's use of information is affected by its cost. Standard economic theory posits that the cost of a given piece of information should not influence the way it is incorporated in belief updating, all else equal. We thus touch upon two known issues concerning individual decision making: the sunk cost fallacy and Bayesian updating. Individuals who are prone to sunk cost effects may behave differently after receiving information for free than after choosing to, or made to, pay for it. Consequentially, sunk costs may have an effect on deviations from Bayesian updating. The extent to which individuals perform Bayesian updating can therefore be exacerbated or alleviated by the cost of information.

To examine these issues we use a laboratory experiment that enables us to control for the selection problems typical of field investigations. That is, we control for the possibility that outcome differences are engendered by the behavior of subjects with different valuation of information rather than by the cost of information itself. We do so by implementing a treatment in which we force subjects to pay for information regardless of their willingness to use it.

We find a significant sunk cost effect on individual belief updating. Subjects who pay for information put higher weight on it relative to subjects who receive identical information at no cost. This effect leads to a shift of updated probability beliefs towards the extremes. Belief updating with costly information can be closer to or further from the Bayesian benchmark compared to belief updating with free information. If subjects under-update their beliefs using free information, then costly information 'pushes' their decision closer to the optimum. In case the opposite occurs, i.e. subjects over-update with free information, then costly information 'pushes' them further away from the optimal outcome. The change in posterior belief is traced to an increase in the extent to which the representativeness bias affects subjects who receive costly information. Cost increases the weight subjects put on newly received information relative to the underlying prior.

These findings have important implications for information acquisition strategies. The effect we describe in this paper should be taken into consideration alongside the price and quality dimensions. Despite the growing availability of free information, tailored information remains a scarce and expensive commodity. Costly information will tend to be of higher quality, and therefore incorporated in decision making to a greater extent than free information. However, the sunk cost effect should reinforce the tendency to overweigh costly information, independent of quality considerations. In practical terms, we might observe situations in which information is used because of sunk cost considerations and not because of how useful it is. Going back to the opening examples of this paper, if the external consultancy firm and the in-house analysts delivered the exact same recommendations, we speculate that the former would be more likely to be adopted because of a sunk cost effect. Analogously, a consumer is more likely to trust a piece of information that was the result of a higher behavioral investment (e.g. time) than an identical piece of information that she obtained through a lower behavioral investment.

We have shown that placing a cost on information can be beneficial or harmful, which suggests that decision makers should consider the implications of providing costly information. This can be beneficial as long as individuals under-update information when it is freely provided. The downside of making information costly is the trade-off between more optimal behavior and lower demand for information. To circumvent lower demand a policy maker can force a cost via a mandatory fee or tax. This fee must be directly related to the information such that an individual does not internalize the fee in advance, thus ignoring its association to the provided information.

Having shown the existence of sunk cost effects on the use of information, further investigations may be of interest. A fundamental issue that should be tackled by future work is the scope and relevance of the price-quality heuristic. In simple terms, costly information might be subconsciously deemed to be of higher quality when in fact it is not. This can lead to, or reinforce, the overweighting pattern we describe. Even though the laboratory provides the ideal conditions for shutting down this heuristic (information is simple, transparent and identical across all treatments), only future work can evaluate its role in the pattern we describe. An interesting extension for a laboratory study is to allow for signals with different informative precision and observe how the sunk cost effect might manifest itself in response. Another natural extension is to explore how updating behavior reacts to differing information costs. For example, if lower costs are found to induce changes in behavior similar to the ones found in this paper, using costly information may also be useful in cases where the demand for information is sensitive to price.

References

Alberoni, F. (1962). Contribution to the study of subjective probability I. *The Journal of General Psychology 66*(2), 241–264.

Ariely, D., U. Gneezy, G. Loewenstein, and N. Mazar (2009). Large stakes and big mistakes. *Review of Economic Studies 76*(2), 451–469.

Arkes, H. R. and C. Blumer (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes 35*(1), 124–140.

Ashraf, N., J. Berry, and J. M. Shapiro (2010, December). Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *The American Economic Review 100*(5), 2383–2413.

Borland, J., L. Lee, and R. D. Macdonald (2011). Escalation effects and the player draft in the AFL. *Labour Economics 18*(3), 371–380.

Buchheit, S. and N. Feltovich (2011). Experimental evidence of a sunk-cost paradox: a study of pricing behavior in Bertrand-Edgeworth duopoly. *International Economic Review 52*(2), 317–347.

Camacho, N., B. Donkers, and S. Stremersch (2011). Predictably non-Bayesian: quantifying salience effects in physician learning about drug quality. *Marketing Science 30*(2), 305–320.

Camerer, C. F. (1987). Do biases in probability judgment matter in markets? Experimental evidence. *The American Economic Review 77*(5), 981–997.

Camerer, C. F. (1995). Individual decision making. In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, pp. 587–703. Princeton University Press.

Camerer, C. F. and R. M. Hogarth (1999). The effects of financial incentives in experiments: a review and capital-labor-production framework. *Journal of Risk and Uncertainty 19*, 7–42.

Camerer, C. F. and R. A. Weber (1999). The econometrics and behavioral economics of escalation of commitment: a re-examination of Staw and Hoang's NBA data. *Journal of Economic Behavior & Organization 39*(1), 59–82.

Charness, G. and U. Gneezy (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry 48*(1), 133–146.

Charness, G., E. Karni, and D. Levin (2007). Individual and group decision making under risk: an experimental study of bayesian updating and violations of first-order stochastic dominance. *Journal of Risk and Uncertainty 35*(2), 129–148.

Cohen, J. and P. Dupas (2010). Free distribution or cost-sharing? Evidence from a randomized Malaria prevention experiment. *The Quarterly Journal of Economics 125*(1), 1–45.

Cunha, Jr, M. and F. Caldieraro (2009). Sunk-cost effects on purely behavioral investments. *Cognitive Science 33*(1), 105–113.

DellaVigna, S. (2009, June). Psychology and economics: Evidence from the field. *Journal of Economic Literature 47*(2), 315–72.

Duh, R. R. and S. Sunder (1986). *Incentives, Learning and Processing of Information in a Market Environment: An Examination of the Base-Rate Fallacy.* Center for Economic and Management Research, University of Oklahoma.

Eger, C. and J. Dickhaut (1982). An examination of the conservative information processing bias in an accounting framework. *Journal of Accounting Research 20*(2), 711–723.

El-Gamal, M. A. and D. M. Grether (1995). Are people Bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association 90*(432), 1137–1145.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*, 171–178.

Friedman, D., K. Pommerenke, R. Lukose, G. Milam, and B. Huberman (2007). Searching for the sunk cost fallacy. *Experimental Economics 10*, 79–104.

Gino, F. (2008). Do we listen to advice just because we paid for it? The impact of advice cost on its use. *Organizational Behavior and Human Decision Processes 107*(2), 234–245.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics 112*(2), 631–645.

Gonzalez, R. and G. Wu (1999). On the shape of the probability weighting function. *Cognitive Psychology 38*(1), 129–166.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics 95*(3), 537–557.

Grether, D. M. (1992). Testing bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization 17*(1), 31–57.

Hammerton, M. (1973). A case of radical probability estimation. *Journal of Experimental Psychology 101*(2), 252.

Herrmann, P. N., D. O. Kundisch, and M. S. Rahman (2015). Beating irrationality: does delegating to IT alleviate the sunk cost effect? *Management Science 61*(4), 831–850.

Ho, T.-H., I. Png, and S. Reza (2015). Sunk cost fallacy in driving the world's costliest cars. Working Paper.

Holt, C. and A. Smith (2009). An update on Bayesian updating. *Journal of Economic Behavior & Organization 69*(2), 125–134.

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

Kahneman, D. and A. Tversky (1973). On the psychology of prediction. *Psychological Review 80*(4), 237–251.

Ketel, N., J. Linde, H. Oosterbeek, and B. van der Klaauw (forthcoming). Tuition fees and sunk-cost effects. *Economic Journal*.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative and methodological challenges. *Behavioral and Brain Sciences 19*, 1–53.

Leuven, E., H. Oosterbeek, J. Sonnemans, and B. Van Der Klaauw (2011). Incentives versus sorting in tournaments: Evidence from a field experiment. *Journal of Labor Economics 29*(3), 637–658.

Offerman, T. and J. Potters (2006). Does auctioning of entry licences induce collusion? An experimental study. *Review of Economic Studies 73*(3), 769–791.

Otto, A. R. (2010). Three attempts to replicate the behavioral sunk-cost effect: A note on Cunha and Caldieraro (2009). *Cognitive Science 34*(8), 1379–1383.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *The Journal of Politics 64*(4), 1021–1044.

Schlag, K. H., J. Tremewan, and J. J. van der Weele (2014). A penny for your thoughts: a survey of methods for eliciting beliefs. *Experimental Economics*, 1–34.

Staw, B. M. and H. Hoang (1995). Sunk costs in the NBA: why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly 40*(3), 474–494.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization 1*(1), 39–60.

Appendix

A Details on the Choice Framework

In this Appendix we provide details on the choice framework and the normative prescriptions (the Bayesian benchmark) for the implemented parameterizations. The family of two-part payoff functions we employ is:

$$F(x,\sigma) = \alpha - \beta \left\| x - s(\sigma) \right\|^{\gamma}$$

where $\sigma \in \Sigma = \{L, R\}$ is the state of the world; $s : \Sigma \to \{l, r\}$ is a state-dependent function such that $s(L) = l$ and $s(R) = r$; $\alpha, \beta, \gamma >> 0$ are parameters; x is the decision maker's choice variable. For $p \equiv \Pr(L)$, expected value maximization yields:

$$x^{*} = \frac{1}{\left(1 + \left(\frac{p}{1-p}\right)^{\frac{1}{\gamma-1}}\right)} \left(r + l\left(\frac{p}{1-p}\right)^{\frac{1}{\gamma-1}}\right) \tag{A.1}$$

which allows us to easily recover p from any observed x.

In our experiment $x \in [0, 100]$, $l = 20$ and $r = 80$ (see Figure 1). As explained in Section 2 there are three possible information conditions, $\phi \in \{Black, White, No\ Draw\}$, which induce different distributions of the lottery. We define a "*Draw*" as "*Black*" or "*White*". Table 6 presents the two parameterizations that were implemented: A was used in the main decision task and B was used in the extra decision task. These parameterizations were chosen such that the loss domain was restricted while providing substantial incentives to Bayesian updating.

	α	β	γ	$\Pr(L)$	$\Pr(Black\|L)$	$\Pr(Black\|R)$	Cost	Exch. Rate
A	6	$9 * 10^{-3}$	1.7	0.4	0.2	0.6	0.3	0.75
B	5.7	$9.25 * 10^{-3}$	1.7	0.7	0.8	0.4	0.25	0.75

Table 6: Parameterizations A and B.

The posterior probabilities $\Pr(L|\phi)$, optimal decisions $x^{*}|\phi$, and the expected values in different information conditions, $E[F(x^{*},\sigma)|\phi]$ and $E[F(x^{*},\sigma)|Draw]$ are provided for both parameterizations in Table 7. Note that the scenarios induce similar expected values across parameterizations, which makes the incentive to optimize and acquire information similar in A and B. The prior probability is different in A and B but the urn composition remains unchanged (it is symmetric).

	$\Pr(L\|\phi)$		$x^*\|\phi$			$E[F(x^*,\sigma)\|\phi]$			$\Pr(Black)$	$E[F(x^*,\sigma)\|Draw]$
	B	W	ND	B	W	ND	B	W		
A	0.18	0.57	58.5	74	44	3.22	4.40	3.15	0.44	3.70
B	0.82	0.44	34	26	55	3.26	4.10	2.76	0.68	3.67

Table 7: Values for Parameterizations A and B. Note: ND, B and W stand for No Draw, Black and White respectively.

B The extra decision task: subject types and a robustness check on selection

As mentioned in Section 2, all subjects go through an extra decision task after completing the main decision task, regardless of treatment. The purpose of the extra decision task is to identify each subject's willingness to pay for, and use, costly information, providing us with a useful characterization of individual heterogeneity and how it can explain selection and behavior.

The extra decision task is based on parametrization B (see Table 7). The idea is to create a decision environment that looks sufficiently different from the main decision task but is equivalent in terms of incentives. In particular, the ratio of the expected gain from using costly information to the expected gain from not using information is similar across A and B. The extra decision task consists of three sequences of ten periods each (see Figure 4 for a diagram). In the first sequence (S_1), information is available for free and in every period. In the second sequence (S_2) information is available at a cost (again at roughly 60% of the expected gain) and in every period. The first two sequences are akin to *Free* and *Costly* with 100% probability of getting information. In the third sequence subjects have to choose between ten periods where they always have to pay for information (which is identical to *Compulsory* with 100% probability of information being available) and ten periods where information is never available.

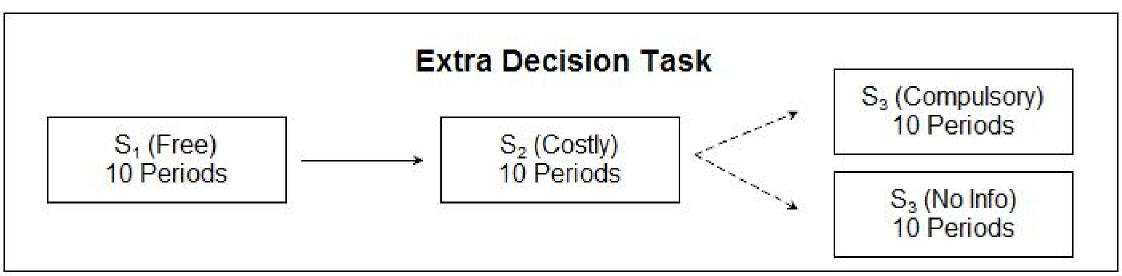

Figure 4: Outline of the extra decision task.

The extra decision task allows us to measure the value of information to subjects, i.e. how their expected benefits compare to the costs they have to incur. We distinguish between two kinds of cost - monetary and cognitive - and propose a classification that takes both into account. Accordingly, a subject buys information if the expected gain of using it (the difference between her expected payoff with and without information) exceeds the cognitive and material costs, denoted by C_C and C_M respectively. In this respect, subjects who observe information incur C_C in S_1 and $C_C + C_M$ in S_2. In S_3 subjects choose whether they want to incur $C_C + C_M$ or not, but their choice is binding for ten periods.

The first two sequences not only provide useful measurements in themselves, they also allow all subjects to experience what it is like to use information for free and at a cost. This is important given that subjects faced different treatment conditions in the main decision task. In fact, we observe no difference in average information use between the different treatments in S_1 and S_3, but we do in S_2. In S_2, subjects in the *Free* treatment are less likely to pay for information than in the *Costly* and *Compulsory* treatments. However, the differences are not significant (MWW, $p \geq 0.12$).

This stylized framework allows us to create an intuitive classification of types. We assume that the expected gain from information exceeds the cognitive costs of processing it if a subject observes information at least 9 out of 10 times in S_1 (i.e. we allow for one mistake); we refer to this as Criterion 1. We further assume that the expected gain from information is higher than its cognitive and material costs when a subject chooses to have information in S_3; we refer to this as Criterion 2. We put forward the following four types of subjects:

- **Type 3** subjects are those who always choose to draw a ball; their expected gain from using information exceeds not only the cognitive cost of using it but also its price in the experiment. Type 3 subjects satisfy both Criterion 1 and Criterion 2.

- **Type 2** subjects are those who choose to draw a ball when it is free but not when it is costly; they expect a net gain from using information but are not willing to buy it at the price charged for it in the experiment. Type 2 subjects satisfy Criterion 1 but not Criterion 2.

- **Type 1** subjects are those who choose to draw a ball neither when it is free nor when it is costly; they do not expect a net gain from using information, even if there is no material cost involved. Type 1 subjects do not satisfy Criterion 1 nor Criterion 2.

- **Type 0** subjects choose to draw a ball when it is costly but not when it is free; they are inconsistent types and they are considered for completeness. Type 0 subjects do not satisfy Criterion 1 but satisfy Criterion 2.

In the experiment, two of the sessions (47 participants, 22 in *Free* and 25 in *Costly*) had a different extra decision task. However, the main decision task was identical across all sessions. The extra decision task was changed after the first two sessions in order to enhance the validity of the type classification. No data from these two sessions is used in the analysis that follows, which is based on the remaining 119 subjects. Moreover, some subjects never drew a ball and thus the number of subjects used is $N = 106$ for a black draw and $N = 103$ for a white draw.

Table 8 presents the distribution of types per treatment, as well as the percentage of information seen in the main decision task by type and treatment. As expected, Type 0 are a residual category in the data. There are no significant differences in the proportion of subject types across treatments. Type also doesn't correlate strongly with observables like risk preferences (Pearson's r=0.18), math courses taken at the university level ($r = 0.12$), math taken in high school ($r = 0.05$) or age ($r = 0.02$). There is no statistical evidence for a difference of observables across types (Pearson's chi-square test $p > 0.13$). This means that our type classification measures the ability and willingness to use information, and not simply mathematical proficiency or risk preferences. The only variable that correlates significantly with type is gender: males tend to be of higher type (Pearson's chi square test $p = 0.02$), albeit moderately ($r = 0.27$). As can be observed, Type 2 and Type 3 subjects observe information more often than Type 1, particularly in the *Free* and *Compulsory* treatments. In the costly treatment the difference between Type 2 and Type 3 subjects goes in the expected direction but falls short of statistical significance (MWW: $p = 0.16$).

	Information acquisition rates			Distribution		
	Free	Costly	Compulsory	Free	Costly	Compulsory
Type 0	0.25	-	0.8	0.02	-	0.03
Type 1	0.24	0.08	0.17	0.14	0.15	0.17
Type 2	0.89	0.52	0.79	0.47	0.43	0.42
Type 3	0.95	0.69	0.93	0.37	0.43	0.39
N	43	40	36	43	40	36

Table 8: Subject types: information acquisition rates in the main decision task and distribution over treatments.

We focus the descriptive analysis on Type 2 and Type 3, as they constitute the majority of our subjects and are the ones for which cost-based selection can take place. Figure 5 presents average posteriors by subject type, treatment and ball draw. Figure 5 shows that Type 2 and Type 3 make similar decisions in *Compulsory* for both draws and in *Free* for the case of a white ball. As expected, the two types make different decisions in *Costly*. Type 3 acquire information more often and therefore tend to exhibit more extreme posteriors. In the case of *Compulsory*, where types cannot choose whether to purchase information, their posteriors are very similar and more extreme than the ones observed in *Free* (with the exception of Type 3 in Black). In other words, both types react similarly when cost is imposed. This pattern is consistent with the sunk cost explanation we suggested. We reinforce our conclusion that information costs have an effect on updating behavior that cannot be explained by selection.

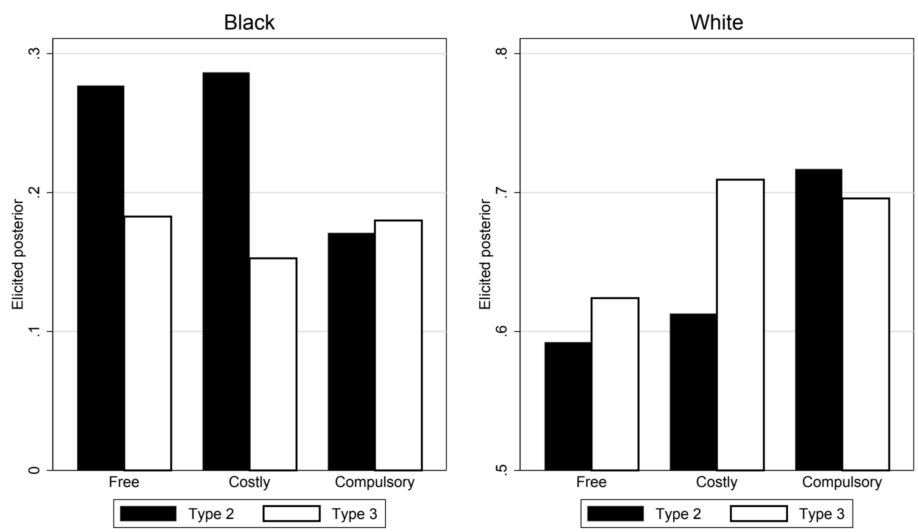

Figure 5: Average elicited probability by subject type.

In order to assess the statistical significance of the relationship between types and behavior we estimate the following model for a black draw and a white draw:

$$\bar{p}_{i\phi} = \alpha + \beta_1 \cdot D_{i\phi}^{Costly} + \beta_2 \cdot D_{i\phi}^{Compulsory} + \beta_3 \cdot D_{i\phi}^{Type\ 1} + \beta_4 \cdot D_{i\phi}^{Type\ 2} + X'_{i\phi} \cdot \gamma + \epsilon_{i\phi} \tag{5}$$

where the dependent variable $\bar{p}_{i\phi}$, and the subscripts i and ϕ, are as defined in equation (1). D^{Costly} and $D^{Compulsory}$ are treatment dummies, and $D^{Type\ 1}$ and $D^{Type\ 2}$ are subject type dummies. The vector of control variables X includes gender, mathematical knowledge and risk aversion. Table 9 presents OLS regression coefficients for posteriors after a black draw and a white draw for different specifications. The baseline category is Type 3 subjects in *Free*. Note that this stratification lowers our sample sizes and results in a lower power of statistical tests. As a result, we will use the 10% level as the threshold for qualitative statements about statistical significance.

The results are consistent with the analysis done in Sub-section 3.1 for the effect of the *Costly* and *Compulsory* treatments. *Costly* and *Compulsory* shift posterior probabilities downwards after a black draw and upwards after a white one, even after controlling for subject types and demographic characteristics, as can be observed in models (3) − (4) and (7) − (8). However, we observe that the inclusion of the Type 2 dummy leads *Costly* to lose significance in the case of Black (model 3). This is the result of *Costly* capturing the behavior of both Type 2 and Type 3 (the other types do not observe much information in this treatment). The behavior of the former can be interpreted as a 'testing the waters' strategy, acquiring information not consistently, which might then be less affected by information costs. Once we account for this interaction effect in model (4), the treatment dummy for *Costly* is again significant.

In no model is the difference between *Costly* and *Compulsory* statistically significant (Wald test, $p > 0.26$). If the difference in mean posteriors between *Free* and *Costly* was driven by those subjects who choose to purchase costly information, then any change between these treatments should be larger than the difference between the *Free* and *Compulsory* treatments. The reason is that the mean posteriors in the *Compulsory* treatment are the weighted average of those subjects who would voluntarily buy costly information and those who would not. As can be seen in Figure 5 this is not the case: the shift from *Free* to *Costly* is comparable to the shift from *Free* to *Compulsory*.

As a last note, the regression results show that some demographic characteristics affect updating behavior. For example, a larger number of math courses leads to a shift towards less extreme posteriors. Gender has a significant effect after a black draw, female subjects tend towards less extreme posteriors, and risk preferences do not have a significant effect. All in all, including both type and demographic variables as controls does not change the results discussed in the main text.

	Black				White			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Costly	-0.07**	-0.08***	-0.04	-0.09**	0.05*	0.07**	0.07*	0.12**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Compulsory	-0.09**	-0.10**	-0.08*	-0.08*	0.10***	0.11***	0.11***	0.11***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Type 1			0.22***	0.22***			-0.20***	-0.20***
			(0.06)	(0.06)			(0.05)	(0.05)
Type 2			0.05	0.02			-0.02	0.01
			(0.03)	(0.04)			(0.03)	(0.04)
Compulsory*Type2				0.11*				-0.11
				(0.06)				(0.07)
Math		0.02***	0.02**	0.02**		-0.02***	-0.02***	-0.02***
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
Female		0.06**	0.07**	0.08**		-0.04	-0.04	-0.04
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
Risk		-0.02	-0.02	-0.02		0.01	0.02	0.02
		(0.02)	(0.02)	(0.02)		(0.01)	(0.02)	(0.02)
Constant	0.27***	0.25***	0.17***	0.18***	0.60***	0.63***	0.66***	0.65***
	(0.02)	(0.05)	(0.06)	(0.06)	(0.02)	(0.04)	(0.06)	(0.06)
N	146	146	106	106	145	145	103	103

Table 9: Regression results: subject types and demographic characteristics. Note: OLS regression. The dependent variable is a subject's mean posterior, $\bar{p}_{i\phi}$. Type 0 subjects are not included in the analysis. Robust standard errors used. ***/**/* indicates significance at the 1%/5%/10% level.

C Experiment Instructions

Below we provide an abridged transcript of the instructions. Square parentheses indicate changes in the sessions with the extra decision task version not analyzed in this paper.

In this experiment you will be asked to make decisions in 70 [80] periods, with one decision per period. The 70 periods are divided in 2 blocks of 40 decisions each. The first block has 40 periods, and the second block has 30 periods. [The 80 periods are divided in 2 blocks of 40 decisions each.] The type of decision is similar, but not identical, across the two blocks. The second block will only start when every participant in this room has finished the first block. You will receive instructions for the second block after the first one is finished. The periods are not timed, which means that you can make decisions at your own pace. We estimate that each block should not take more than 40 minutes to complete.

Your earnings will be determined according to your performance in the experiment. Out of each block, 3 periods will be randomly selected to be paid (that is, 6 periods in total). All payoffs in the experiment are expressed in tokens. Each token in the experiment is worth 0.75 Euro.

First Block:In each period you can be in one of two States, Left and Right. There is some probability that you are in Left and some probability that you are in Right. Think of this as tomorrow's weather in Sydney: with a certain probability tomorrow will be cloudy and with a certain probability tomorrow will be sunny, but we don't know for sure what the weather in Sydney will be tomorrow. The same applies to the States in this experiment. The probability that the state is Left is 40% and the probability that the state is Right is 60%. As you can see, the two probabilities sum to 100%. These probabilities will be shown on your screen at all times.

Your decision in each period is to pick a number from 1 to 100. You can pick numbers in steps of 0.5, which means that 24 and 24.5 are possible, but 24.4 and 24.6 are not. Your payoff in each period will depend on your decision (the number you choose) and the actual State (Left or Right). Below you can see two graphs showing how the payoffs depend on your decision and the State:

(a graph similar to the one in Figure 1 was shown here)

These graphs show that if the State is Left, choosing 20 yields the highest payoff, and if the State is Right choosing 80 yields the highest payoff. However, if 20 is chosen and the State is Right, a negative payoff results. The same is true if 80 is chosen and the state is Left. Given that the actual state is not known when you must make your decision, choosing other values can make sense.

You can find a Table with the payoffs for all possible combinations of decisions and States in the last sheet. You will also be able to see those payoffs on the computer screen before making your decision.

In each period, a basket with 5 balls is presented. Some balls are black and some are white. The composition of the basket depends on the State. If the state is Left then there is one black ball and four white balls in the basket. If the state is Right then there are three black balls and two white balls in the basket.

(a graph depicting the distribution presented in Table 1 was shown here)

In each period, there is a 50% chance that you can see a ball drawn from the basket. Note that when the ball is drawn you still do not know what the State is, which means that you don't know from which basket composition you are drawing the ball.

To summarize, the events in each period of the first block occur in the following order:

1. The State is randomly determined. You do not know what the State is at this point.
2. With a 50% chance you have the option of seeing a ball drawn from the basket.
3. You make your decision.
4. The State is revealed and your payoff is known.

Second Block**:** You will now begin the second block of the experiment. Note that the State probabilities and the payoffs have changed from the first block you have just finished.

In this block the probability that the state is Left is 70% and the probability that the state is Right is 30%. As you can see, the two probabilities sum to 100%. These probabilities will be shown on your screen at all times.

Your payoff in each period will depend on your decision (the number you choose) and the actual State (Left or Right). Below you can see two graphs showing how the payoffs depend on your decision and the State:

(a graph similar to the one in Figure 1 was shown here)

These graphs show that if the State is Left, choosing 20 yields the highest payoff, and if the State is Right choosing 80 yields the highest payoff. However, if 20 is chosen and the State is Right, a negative payoff results. The same is true if 80 is chosen and the state is Left. Given that the actual state is not known when you must make your decision, choosing other values can make sense.

You can find a Table with the payoffs for all possible combinations of decisions and States in the last sheet. You will also be able to see those payoffs on the computer screen before making your decision.

In each period, a basket with 5 balls is presented. Some balls are black and some are white. The composition of the basket depends on the State. If the state is Left then there are four black balls and one white ball in the basket. If the state is Right then there are two black balls and three white balls in the basket.

(a graph depicting the distribution presented in Table 1 was shown here)

Note that when the ball is drawn you still do not know what the State is, which means that you don't know from which basket composition you are drawing the ball.

This block is composed of three sets of 10 decisions. Each set differs in the manner in which a ball can be drawn from the basket. Further instructions will be given on the computer screen before each set of 10 decisions. [In each period, there is a 50% chance that you can see a ball drawn from the basket.]