TI 2012-143/VII
Tinbergen Institute Discussion Paper

Information at a Cost

Pedro Robalo[1,2]
Rei Sayag[3,2]

[1] *Faculty of Economics and Business, University of Amsterdam;*
[2] *Tinbergen Institute;*
[3] *Erasmus School of Economics, Erasmus University Rotterdam.*

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Information at a cost

Pedro Robalo
University of Amsterdam
Tinbergen Institute

Rei Sayag*
Erasmus University Rotterdam
Tinbergen Institute

October 2012

Abstract

The supposed irrelevance of historical costs for rational decision making has been the subject of much interest in the economic literature. In this paper we explore whether individual decision making under risk is affected by the cost of the supplied information. Outside of the lab, it is difficult to disentangle the effect of the cost of information itself from the effect of self-selection by individuals who tend to gain the most from this information. We thus create an environment in the lab where subjects are offered additional, useful and identical information on the state of the world across treatments. By varying the cost of information we can distinguish between selection and sunk cost effects. We find a systematic effect of sunk costs on the manner in which subjects update their beliefs on the state of the world. Subjects over-weigh costly information relatively to free information, which results in a 'push' of beliefs towards the extremes. This shift does not necessarily lead to behavior more attuned with Bayesian updating.

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

JEL codes: C91; D81; D83.

*For their useful comments, we would like to thank Johannes Abeler, José Apesteguia, Christoph Brunner, Gary Charness, Dan Levin, John List, Ariel Rubinstein, Sally Sadoff, and seminar participants at the University of Amsterdam; Erasmus University Rotterdam; the Tinbergen Institute, Amsterdam; MBEES 2012, University of Maastricht; CCC 2012, University of East Anglia; the Economic Science Association Conference 2012, New York; and the Choice Lab, Norwegian School of Economics. Financial support by the Priority Research Area Behavioral Economics of the University of Amsterdam and the Erasmus University Trustfonds is gratefully acknowledged. Robalo acknowledges financial support from Fundação para a Ciência e Tecnologia through Grant SFRH/BD/43843/2008. Sayag (corresponding author): Erasmus School of Economics, Burg. Oudlaan 50, 3062 PA Rotterdam, the Netherlands . E-mail: sayag@ese.eur.nl. Robalo: CREED, Amsterdam School of Economics, Roetersstraat 11, 1018 WB Amsterdam, the Netherlands. E-mail: p.robalo@uva.nl.

1 Introduction

Information plays a crucial role in supporting decisions made by individuals and organizations. This paper investigates whether the cost paid for information influences the way it is used in decision making. As an illustration, consider a region at risk of pandemic. Assume that a simple and effective prevention method exists (washing your hands twice a day, for example). The local authorities simply have to make sure that knowledge of this prevention method reaches individual citizens, i.e. they have to make citizens aware of this information (through the distribution of a leaflet, for example). Assuming that the cost of doing so is negligible compared to the welfare gains, should the authorities distribute the information for free or charge for it? Conditional on receiving the information, the behavior of a rational individual should not depend on the price paid for it. However, we conjecture otherwise: decision makers might put a higher weight on information they had to pay for, and paying for information can interact with optimization behavior. If this is true, a possible policy implication is that the cost of information can be used to steer decision makers' behavior in the right direction.

Underlying our conjecture is the possibility that individuals fall prey to a variant of the sunk cost effect (Thaler 1980), and 'use' information relatively more when it comes at a cost.[1] We contribute to the accumulated evidence on the sunk cost fallacy by analyzing it from a new perspective, namely by investigating whether it exists with respect to acquired information in a scenario of decision making under risk. If a relationship exists a follow-up question is how it affects decision making, namely 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 be different along several dimensions from individuals who choose not to do so. The laboratory allows us to correct for these selection issues through carefully constructed procedures. One way in which we disentangle selection from sunk cost effects is by imposing the cost of information on subjects. This is something that is easily done in the lab, but arguably difficult to implement in the field.[2] Moreover, the lab allows us to assess the extent to which individuals value information and are able to use it; in other words, we can identify different types of individuals concerning their revealed demand for information.

The sunk cost fallacy's main trust is that only marginal costs and benefits should matter for decision-making. The vintage normative prescriptions ("don't push yourself through a movie which you are not enjoying", for example) are one of the first lessons that business and economics students are exposed to. And indeed the sunk cost fallacy still seems to plague many courses

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

[2] Field tests of sunk cost effects in product use have been carried out (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, measuring product usage (for the cited works, a theater season ticket, a bottle of water disinfectant and bed nets, respectively) is easier than measuring information usage.

of action, be it continuing a failed relationship because one already invested many years in it or a failure to withdraw from a lost war because of an extensive death toll. Thaler put forward a compelling rationale for why people fall prey to the sunk cost fallacy based on loss aversion. Given the convexity of losses, a decision-maker facing a risky investment has an incentive to recover an incurred loss since the increase in utility of a gain will be larger than what a further comparable loss would entail. Despite the abundance of casual and anecdotal evidence, the literature's verdict on the sunk cost fallacy is more mixed. The pioneering field experiment of Arkes and Blumer (1985) found that granting a random discount for a theater season ticket significantly decreases attendance. 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. The experimental laboratory evidence is slightly more supportive of the sunk cost fallacy. 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 occasionally at best. However, in an Industrial Organization setting, both Offerman and Potters (2006) and Buccheit 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 purely 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. An attempt at replicating these findings was not successful (Otto 2010). Gino's contribution (2008) is methodologically close to our work, but focuses on the role of the costs of advice: it is shown that the (exogenously determined) cost of another subject's advice influences its use. Subjects who were exposed to paid advice incorporated it significantly more in their decisions than those who obtained it for free. From the competing explanations, the author shows that sunk cost effects drive the results.

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 an overview). 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 popular heuristics. To be sure, the verdict on whether "man is a Bayesian" is still out. When combining information on prior probabilities of the possible states of the world with informative state-dependent signals, three main inter-related phenomena are observed (see Camerer 1995 for a detailed overview). First, individuals often exhibit conservatism in their choices, failing to use the signal to the extent normatively prescribed by Bayes' formula (e.g. Eger and Dickhaut 1982). Second, there is a systematic tendency for subjects to overweigh the signal in their judgment relative to prior probabilities, often referred to as the 'base rate fallacy' (see Koehler 1996 for an appraisal of the literature). 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 the signal in a given state, he will tend to overweigh the probability that this state will occur (often referred to as 'exact representativeness'). Early evidence (e.g. Kahneman and Tversky, 1972 and 1973) showed that representativeness was a serious and systematic bias, leading these authors to claim that "man is not a Bayesian at all" (1973). 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, the "most likely rule that people use is 'Bayes's rule.'" (El-Gamal and Grether 1995), and only then representativeness. Experimental market tests of this heuristic (Duh and Sunder 1986 and Camerer 1987) have shown that behavior converges to Bayesian and the observed deviation is mostly explained by representativeness. In sum, with respect to conservatism, the base rate fallacy 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). Not only that, base rates' "degree of use depends on task representation and structure" (Koehler 1996). Building upon these conclusions, we ask a natural question: can the cost of information influence the extent to which conservatism, the base rate fallacy and representativeness prey on decision makers? In other words, can the cost of information mediate the difficulties posed by Bayesian updating (as emphasized by economists) and a tendency to disregard underlying prior probabilities (as documented by psychologists)? If that is the case, the cost of information can be used to dampen some of the shortcomings associated with decision making under risk. It is important to note that each setting poses its challenges and therefore general prescriptions are probably useless; nevertheless, an existence result would constitute a pre-requisite for further investigations into context-specific fine-tuning where information cost is the control variable.

In our design, each participant has to make a number of discrete decisions with state-dependent payoff consequences. There are two states with known and constant priors. Subjects sometimes have the opportunity of reducing uncertainty by drawing a sample (a 'ball') from a state-dependent lottery (an 'urn' with balls). Our treatments change the way in which this information is made available: in the Free treatment it is made available at no cost, while in Costly it has to be

[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."

purchased. A treatment where the cost is imposed on subjects (Forced) corrects for selection while leaving the role of cost intact. Moreover, and for all treatments, subjects subsequently go through a reduced version of the two first treatments and are asked to make a decision which allows us to distinguish between individuals who are willing to purchase costly information and those who are not. We observe subjects' revealed demand for information and classify them by types, taking into account both the cognitive and the material costs of information. This allows us to further analyze the role of selection in our results.

Our results show that individual decisions are in line with the described biases, with deviations from the Bayesian normative model explained both by under- and over-updating. Paying for information leads to an over-weighting of newly obtained information, which leads to more extreme moves in the posterior in the Costly and Forced treatments as compared to 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 the data shows no significant differences between Forced and Costly. Moreover, subject types do not explain the overall pattern, reinforcing the sunk cost explanation. Regarding choice optimality, more extreme choices can lead to better or worse decisions. Most subjects benefit from having access to information (regardless of the cost) as it allows them to reduce uncertainty. However, some subjects do not benefit from information as the return derived from reduced uncertainty does not compensate for the cost paid for information. Our main conclusions is thus that costly information will be assigned a higher weight in decision making under risk but this does not always lead to more optimal behavior. From a policy perspective, charging for information is beneficial if the decisions made using free information correspond to a situation of Bayesian under-updating, since costly information leads subjects to put a higher weight on newly obtained information. With respect to the example mentioned in the beginning of this section, the authorities could charge for information if they realize that individual decisions do not incorporate the information content of the flyer to the desired extent. The paper is organized as follows. Section 2 presents the experimental design, Section 3 presents our results and Section 4 presents a small exercise on information pricing. A final section concludes.

2 Experimental Design

Each subject has to make decisions in two blocks: the Decision block and the Identification block, comprising 40 and 30 periods, respectively. Our analysis focuses on the data obtained from the Decision block, while data from the Identification block is used to account for the discussed selection issues. The decision was identical across the two blocks except for parameterization. Paper instructions were distributed in the beginning of each block, which subjects were asked to read silently. Each block started after all subjects had finished reading the instructions. A set of practice questions to test understanding of the experiment were administered before the start of the

Decision block. In the experiment all values are expressed in tokens, which were converted at an exchange rate of 0.75 Euro per token. Subjects were paid for six randomly determined periods, three from each block.

2.1 Choice Framework

We presented subjects with an intuitive, yet non-trivial individual choice task in which information can be used in a Bayesian fashion. In each period the decision maker faces one of two states of the world (**L**eft and **R**ight), for which probabilities are known: $p \equiv \Pr(L)$. In the Decision block $p = 0.4$.[4] The payoffs are determined by a state-dependent scoring function (see Figure 1).[5] The parameterizations were chosen such that the loss domain was restricted while still providing substantial incentives to perform Bayesian updating.

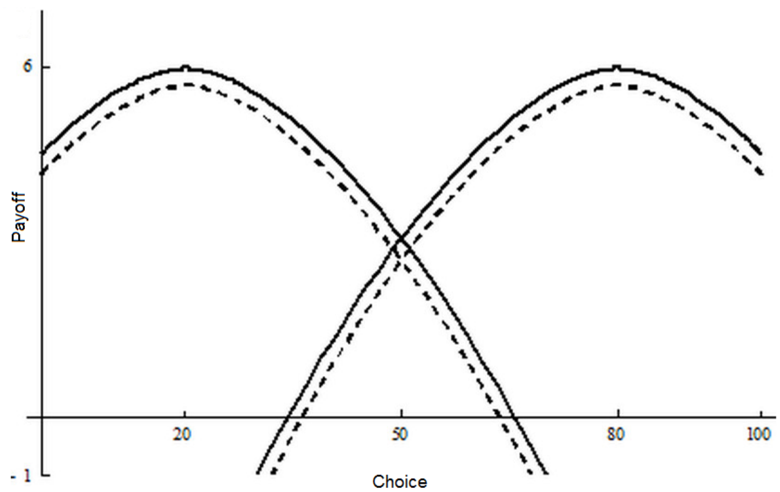


Figure 1: The scoring function. Note: the solid (dashed) line corresponds to the Decision (Identification) block.

Subjects choose a number between 0 and 100 in steps of 0.5. If the state is L (R) the optimal choice is 20 (80). Choices below 20 and above 80 are strictly dominated. The information on the two state-dependent payoffs is made available to subjects in three distinct ways: on the screen (updated every time the subject adjusts her choice before making it final), in graphical format and in table format (both in the paper instructions). Our state-dependent payoff function is an adjusted quadratic scoring rule, which was preferred to proper scoring rules that are robust to

[4]We chose not to implement symmetric priors for a two reasons: first, the task could become trivial (Camerer 1987) or invite the usage of "obvious" (but possibly wrong) heuristics; and second, it would make the alignment of incentives and moves in the posterior across blocks impossible to achieve. The second aspect is important because we want to identify types in an environment (the Identification block) that is as similar as possible to the Decision block.

[5]See Appendix A for a detailed description of the choice environment and derivation of optimal decisions according to the normative model. See Appendix D for a snapshot of the experiment (practice questions and main decision screen).

probability sophistication (e.g. Offerman et al. 2009) for two reasons. First, in our case it is not problematic if risk attitudes bias subjects' decisions in some direction as we are looking for treatment effects (and we further control for risk attitudes statistically). Second, and importantly, the traditional scoring rules do not provide a substantial incentive to update beliefs unless radical moves in the posterior are observed. In other words, we need a scoring rule that is steep enough in the region where probability updating takes place.

The information signal we provide to subjects is a lottery: an "urn" filled with balls. There are five balls in the urn, some black and some white. The distribution of balls is itself state-dependent, but does not change across periods within a block and is visible to the decision-maker before every draw. In our design, drawing a ball from the urn is informative of the state of the world, i.e. the probability of the realized state being L or R should be updated after drawing a ball from the urn. In the Decision block, the urn contains one (three) black balls if the state is L (R) (see Figure 2).

	Left	Right
Prior	0.4	0.6
Urn	■□□□□	■■■□□

Figure 2 - Decision block: prior and lottery distribution (urn) in the two states of the world.

2.2 Treatments

We implement our treatment conditions by varying the way in which the information is made available to subjects in the Decision block. The Identification block is identical across treatments. In Free the ball can be drawn from the urn at no cost. In Costly a ball can be drawn at a cost. In Forced the price of drawing a ball is imposed upon subjects (subjects are told that "a ball has been drawn for them"). In Costly a subject buying a ball observes it automatically while in Free and Forced subjects can choose whether to see the drawn ball or not. In the Decision block there is a 50% chance that subjects can draw a ball from the urn in each period of Free and Costly, or a ball is drawn for them in the case of Forced. If information was available in all periods we would run the risk of subjects automatically discounting the costs of information to be incurred in the beginning of the experiment, which would dissolve the psychological impact of having the cost imposed on them. This also forces subjects to experience decisions without information, which provides us with individual decisions made without an information signal - a likely anchor for decisions when information is made available. In the Costly and Forced treatments the information is priced at $c = 0.3$ tokens, which is roughly 60% of the expected gain if expected utility maximization with Bayesian updating is performed by a risk- and loss-neutral decision maker.

The Identification block uses the same framework with a slightly different parameterization and a similar (reverse) prior, $\Pr(R) = 0.3$. The idea is to create a decision environment that is equivalent in terms of incentives but that looks sufficiently different for it not to be trivial nor invite the application of the decision rules employed or learned in the Decision block. In particular, the ratio of the expected gain from using costly information to the expected gain from not using information is similar across blocks (see Appendix A for details).

The Identification block consists of three sequences of ten periods each. Information is available in every period. In the first sequence (I_1), information is available for free. In the second sequence (I_2) information is available at a cost (c =0.25 tokens, which is again roughly 60% of the expected gain). The first two sequences are akin to the Free and Costly treatments with a 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 Forced with a 100% probability of having information) and ten periods where information is never available. See Figure 2 for a time-flow diagram of the experiment.

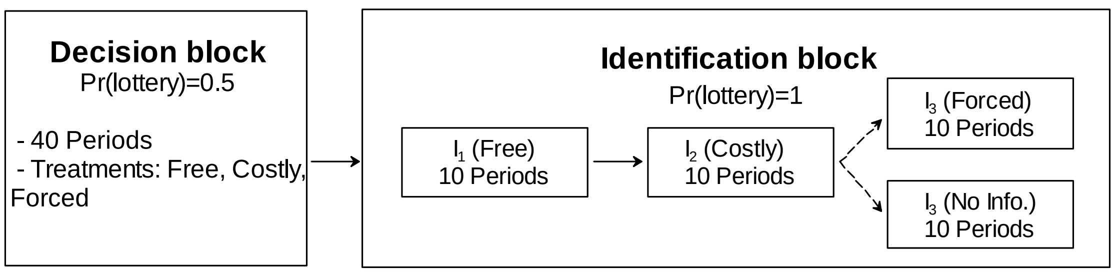


Figure 2: Outline of the experimental design.

The Identification block 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 can distinguish between two types of cost: monetary and cognitive. A subject buys information if:

$$V_i\left(Draw\right) - C_{1,i} - C_2\left(\theta\right) \geq V_i\left(No\ Draw\right)$$

where $\theta \in \{Free, Coslty, Forced\}$, $V_i\left(.\right)$ is the expected payoff of a subject (which depends on many cognitive factors like aptitude, mathematical training, confidence, etc.), $C_{1,i}$ is the cognitive cost of processing information, and $C_2\left(.\right)$ is the monetary cost of information acquisition (equal to 0 in I_1 and equal to c in I_2 and I_3).[6] In this sense, in exchange for information, subjects incur C_1 in I_1 and $C_1 + C_2$ in I_2. Subjects make this choice in every period in I_1 and I_2. In I_3 subjects also choose whether they want to incur $C_1 + C_2$ or not, but their choice is binding for ten periods. This stylized framework allows us to create an intuitive classification of types.

A subject who chooses not to see information in I_1 considers the cognitive cost of processing it superior to the benefits. A subject who chooses not to buy information in I_2 finds the sum of the cognitive and material costs of information higher than the benefit. Sequence I_3 measures the same relationship, but the choice is presented in a dichotomous way. The first two sequences not

[6] See Appendix A for further details.

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, especially considering that they face different treatment conditions in the Decision block.[7] Combining data from sequences I_1 and I_3 allows us to classify subjects in a way that improves our understanding of the major selection issues at hand. In particular, we classify subjects into four types:[8]

	$\Delta V_i(.) \leq C_{1,i} + c$	$\Delta V_i(.) > C_{1,i} + c$
$\Delta V_i(.) \leq C_{1,i}$	Type 1	Type 0
$\Delta V_i(.) > C_{1,i}$	Type 2	Type 3

Table 1: Subject types.

Type 3 individuals are those whose expected gain from using information exceeds not only the cognitive cost of using it but also the monetary cost charged for it. Type 2 individuals expect a net gain from using information but are not willing to buy it at price c. Type 1, on the other hand, do not expect a net gain from using information, even if there is no material cost involved. Type 0 are inconsistent types and they are considered for completeness (as we will see, they are a residual category in the data). We assume that $\Delta V_i(.) \leq C_{1,i}$ if a subject observes information less than 9 out of 10 times in I_1, and $\Delta V_i(.) \leq C_{1,i} + c$ if a subject chooses to have no information in I_3 (in Section 3.3 we analyze the distribution of types that we obtain in light of these criteria).[9]

In order to control for risk attitudes and demographic characteristics in our 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.[10]

3 Experimental Results

The experimental sessions were run at the CREED laboratory of the University of Amsterdam between February and May 2012; they were programmed and conducted with the experiment software z-Tree (Fischbacher 2007). A total of 166 subjects participated in 8 sessions, recruited

[7] We observe no difference in average information use between the different treatments in I_1 and I_3, but we do in I_2. In I_2, subjects in the Free treatment are less likely to pay for information than in the Costly and Forced treatments. This however is only marginally significant for the difference between Free and Forced (Mann-Whitney-Wilcoxon rank-sum test, p=0.11) and it is not significant between Free and Forced.

[8] Where:

$$\Delta V_i(.) = V_i(Draw) - V_i(No\ Draw)$$

[9] The first criterion is employed as we are looking for subjects who would buy information whenever it is free (which is 10 times in I_1) while allowing for one mistake.

[10] The risk attitute 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. In terms of statistical information, the questionnaire asked whether subjects had had Math in high school, how many Math courses they had completed at university, as well as their gender, age, and major.

online from a subject pool of students at the University of Amsterdam. Fifty-five per cent of the participants were male and 57% were Business or Economics majors. The typical session took 1 hour and 20 minutes with average earnings of 24 Euro (which includes a show-up fee of 7 Euro). Two of the sessions (47 participants, 22 in Free and 25 in Costly) had a different Identification block.[11] Unless mentioned otherwise, all data discussed in this section pertains to decision making in the Decision block. Sub-section 3.1 describes the data. Sub-section 3.2 analyzes the difference in decision making across treatments. Sub-section 3.3 expands the analysis by including subject type data.

3.1 Data Description

Table 2 provides a summary of descriptive statistics for the collected data. Differences in individual traits are not statistically significant across treatments. Average period payoff is significantly different between the Free treatment and the Forced treatment. This is to be expected as subjects incur no costs in the Free treatments as opposed to the Costly and Forced treatments. Percentage of information seen refers to the fraction of times subjects choose to observe information when it is available. Naturally, when information is costly and optional, less subjects choose to observe it. The Costly treatment is thus significantly different from the Free and Forced treatments in this respect. We observe that subjects choose not to see information (draw a ball) sometimes, even when it is free or already paid for. This is possible as in all treatments we let subjects have the option of not drawing a ball, and reasoned in terms of the stylized model discussed in Section 2.2. That is, some subjects, denoted as Type 0 and Type 1, find the cognitive costs of using information higher than the benefits. Additionally, many subjects experiment with drawing and not drawing a ball and thus don't observe information in some of the periods.[12]

	Free	Costly	Forced
N	65	65	36
Risk	2.06	1.93	2.10
Math courses	2.38	2.95	2.55
% Female	37%	48%	44%
Average period payoff	3.10	3.00	2.84
% Information seen	79%	56%	74%

Table 2: Summary statistics

For ease of exposition and later analysis we define benchmark decision making as the optimal

[11] The Decision Block was identical across all sessions. The Identification Block was changed in order to enhance the validity of the type dichotomy. For this reason no data from these two sessions is used in analyses containing type variables.

[12] Ignoring type 0 and 1 subjects, % information seen changes to 92% in the Free treatment and 87% in the Forced treatment. Of all subjects, only 6% in Free and 11% in Forced chose never to draw a ball (The figure is 21% in Costly).

choices made by a rational, risk- and loss-neutral individual (hereafter, the Bayesian benchmark). The Bayesian benchmark decisions are 74, 44 and 58.5 while the data averages are 69.4, 39 and 55.6 for a Black draw, a White draw and No draw, respectively. Note that our interest lies mainly in the difference in decision making across treatments and not in individual deviations from the theoretical Bayesian benchmark. Nevertheless, we briefly elaborate on the possible explanation for the deviation of average decision making from the benchmark.

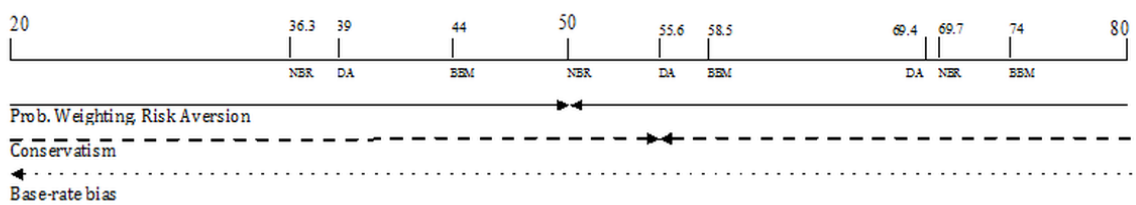


Figure 3: Effect of biases on decision making. Note: NBR: No Base Rate - individual choice with a full base-rate bias. DA: Decision Average - the data mean decision. BBM: Bayesian Benchmark.

These deviations are all statistically significant[13], and are in line with several documented behavioral biases in decision-making under risk (Camerer 1995, Gerther 1992), namely conservatism, probability weighting and base-rate bias. Risk aversion can also be a cause of deviation. Figure 3 graphically illustrates the effects of these biases and risk aversion on decision making. Conservatism pertains to individuals' tendency to underweigh new information, biasing their choice towards the prior. After a ball draw this bias should lead to decisions which are closer to the No draw average decision of 55.6 than the benchmark. Probability weighting describes a tendency to overweigh low probabilities and underweigh high probabilities, increasingly with more extreme probabilities (Holt and Smith 2009). In our design this phenomenon is likely to have a strong effect after a Black ball. The probability that the state of the world is Left after such a draw, using Bayesian updating, is 0.18. Probability weighting causes an individual to asses that the likelihood of Left is higher than this. This bias results in a deviation towards 50 after both draws, similarly to the effect expected with conservatism. The base-rate bias pertains to the tendency of individuals to underweigh the prior when receiving new informations. That is, to update their belief as if the prior is closer to 0.5 than it is in reality. Note that this is *not* the opposite of conservatism. 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 a White draw and under-update it after a Black draw. In case an individual exhibits base rate bias she thus deviates towards 20 after any ball draw. Risk averse subjects are expected to shift their decisions towards less risky options, that is to deviate

[13] All differences are at the 0.00 significance level (mean comparison t-test)

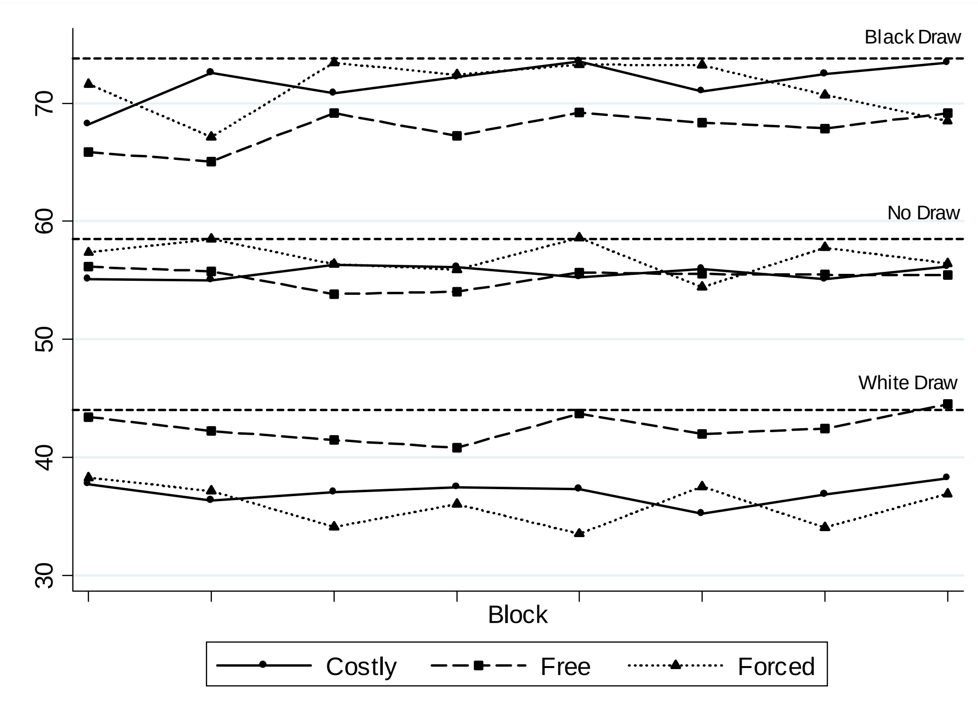


Figure 4: Average decisions by periods. Note: Decision averages for each block are calculated using all decisions made in a 5 period interval (8 blocks of 5 periods) by all subjects within an information type. For example, the top solid line oscillating around 70 represents the 5 period average of subjects' decisions who drew a Black ball in the Costly treatment. The dashed straight lines denote the Bayesian benchmark.

towards 50. With no Bayesian updating, risk aversion and probability weighting are the only two described effects that can influence decisions without a ball draw. Both of them can explain the deviation in average decision making with no additional information. As no single bias can explain the deviation after a White draw and a Black draw, only some combination of the discussed biases, and risk aversion, can explain it.[14]

3.2 Treatment effects

We now begin with an analysis of the aggregate treatment outcomes. Figure 4 presents five-period average decisions over the duration of the Decision block by treatment and information condition (No draw, Black draw and White draw). Decision averages visibly differ across treatments and in all periods after a White draw and a Black draw. No such difference is discernible after No draw. Table 3 shows the aggregate treatment averages by information condition and treatment. After both a White draw and a Black draw there are significant differences between the Costly and

[14]Base-rate bias alone can not justify this deviation. If base rate bias was complete (i.e. the perceived prior is $p = 0.5$) the optimal decision after a Black draw would be 69.5, slightly higher than average decision. After a White draw it would be 36, below the actual deviation. A combination of base-rate bias with conservatism and probability weighting is more likely.

Free treatments (two-sided Mann-Whitney-Wilcoxon rank-sum test, MWW hereafter: $p = 0.02$ and $p = 0.01$, respectively), and between the Forced and Free treatments (MWW: $p = 0.04$ and $p = 0.01$, respectively). No significant differences are found between the Costly and Forced treatments. Without a draw from the urn there are no significant differences between any of the treatments. The shift in average decision making between both the Costly and Forced treatments and the Free treatment is thus significant only after a ball draw. It is an upward shift after a Black draw and downward one following a White draw.

	A Black draw	A White draw	No draw
Free	66.81	41.84	54.93
	(1.32)	(1.39)	(0.77)
Costly	71.23	37.99	56.15
	(1.44)	(1.61)	(0.85)
Forced	71.39	35.31	56.01
	(1.48)	(1.78)	(1.34)

Table 3: Decision averages by treatment. Note: Standard errors in brackets.

Figure 5 presents the cumulative distribution functions of individual decision by information condition. A decision of subject i after observing information condition ϕ is defined as the average of all of her decisions after observing information condition ϕ from period $t = 11$ onwards:[15]

$$\bar{d}_{i\phi} = \frac{\sum_{t=11}^{40} d_{i\phi t}}{n_{i\phi}} \tag{1}$$

As with Figure 4, a shift in decision making between the Free treatment and the Costly and Forced treatments is clearly seen after a ball is drawn. Decision distributions in Costly and Forced first-order stochastically dominate the decision distribution in Free after a Black draw, and are first-order stochastically dominated after a White draw. The differences in distributions are all significant at the 5% level (two-sample Kolmogorov-Smirnov test) except for the one between the Forced and the Free treatments after a Black draw.[16]

From Figure 4 and Figure 5 it is clear that the availability of costly information 'pushes' subjects' decisions more to the extremes. Subjects who incur a cost tend to make more extreme choices vis-a-vis those who did not. Thus, as subjects who pay for information behave differently than those who do not, we conclude that a sunk cost effect exists in our experiment. Thaler's (1980) classic explanation of the sunk cost phenomenon describes the observed data well. A

[15] The choice to discard the first 10 periods was made in order to lower the effect of learning. The results are robust for the inclusion of all of 40 periods. This definition of decision average does not include variations is the number of times an information type was received over subjects. This potential effect is accounted for by subject types in Sub-section 3.3. A small minority of individuals (6%) have a strictly dominated average decision outside the range $[20, 80]$. Only 1% are outside the range $[19, 81]$.

[16] See Table 10 in the Appendix for the p-values of the Kolmogorov-Smirnov test for equality of distribution functions.

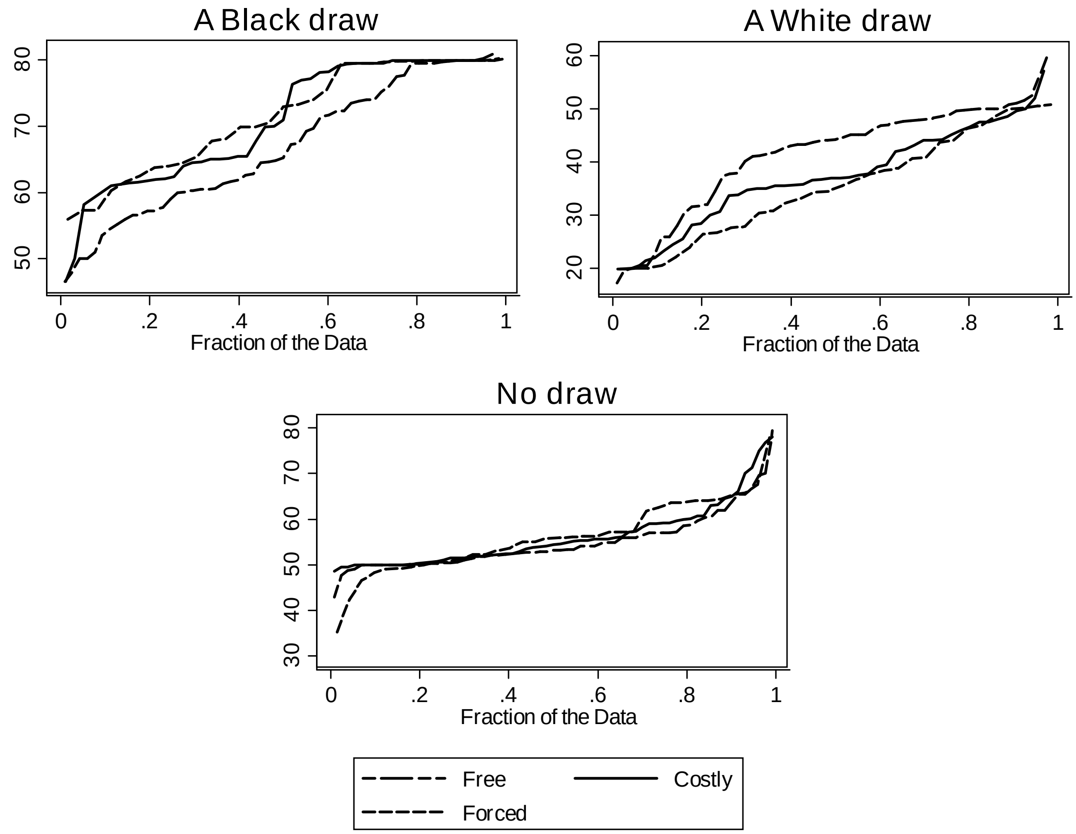


Figure 5: Distributions of individual decision making. Note: Outliers were discarded (one observation after a black draw and one after a white draw).

subject paying for information, found in a loss state, has a higher marginal benefit of payoff 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 levels may change the importance behavioral biases play in the use of information relatively to lower efforts. Another manner in which sunk cost can lead to this change is by increasing the attention a subject places on the ball draw she paid for, and consequentially the relative salience of the information obtained. This would be comparable to drawing bright red circles around the new information after a draw.

Any explanation for the observed data must motivate the change in decision making towards the extremes in the Costly and Forced treatment relatively to the Free treatment. Both higher effort and an increase in saliency can influence the effect behavioral biases have on individual decision making. An increase in the strength of the base-rate bias due to costly information is a natural consideration. A subject who focuses more on the new information she just paid for discounts the underlying prior. This explanation fits the observed behavior after a White draw but not after a Black draw. If the base-rate bias increases with a costly ball draw, average decision after a Black draw should shift towards 50 relatively to a free ball draw while we observe the opposite. Representativeness bias may describe the data well. This bias generally states that if a drawn sample matches one of a number of possible populations the most, the assessed probability of that population will be higher than Bayesian updating dictates. 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 if Left and 2 are white if Right), and a Black ball to be more representative of Right. This would directly lead to more extreme decisions. Another likely possibility is diminishing conservatism in updating of new information. In this case the change in decision making should indeed be away from 50 and towards the extremes. Risk loving behavior resulting from convex utility function in a loss-frame can also explain the observed shift in decision making. We find this explanation unlikely though as the cost of information is much too small to reasonably explain the observed change in behavior. In sum, we identify a number of possible channels through which sunk cost effects can operate, although we are not able to single one out.

Result 1: *Paying for information alters individual decision making in a systematic manner. After incurring a cost individuals overweigh the newly acquired information. This behavior can be explained by a sunk cost effect.*

The difference between the Costly and Free treatments for decisions with a draw from the urn could possibly be explained by selection. That is, the two samples are not identical as subjects who choose to pay for information make different decisions than those who are only willing to observe information at no cost. The average decision 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 do not. The average decision in the Costly treatment after a ball draw is thus the outcome of only one sub-group's decisions. Selection does not play a part in the Free and the Forced treatments, as evidenced by the similar information acquisition rates in Table 2. Any difference in decision making should be thus attributed to the difference in cost incurred by the subjects between these two treatments. Selection effects are further discussed in the following sub-section.

We now turn to decision optimality across treatments, where we define optimality as the absolute distance from the Bayesian benchmark. Table 4 presents the treatment means by information type. The average decision in the Free treatment is the most optimal decision after a White draw, but the least optimal after a Black draw.

	A Black draw	A White draw	No draw
Benchmark	74	44	58.5
Free	66.8	41.8	54.9
Costly	71.2	38	56.1
Forced	71.4	35.3	56

Table 4: Treatment averages

Depending on the true 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, as 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.

Result 2: *Costly information does not necessarily lead to a more optimal use of information.*

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 results other than better performance (Camerer and Hogarth 1999, Ariely et al. 2009, Leuven et al. 2011).

3.3 Selection effects

We now extend the analysis to include type variables as described in Sub-section 2.2 and summarized in Table 1.[17] Table 5 presents summary statistics for these subjects. There are no significant differences in the proportion of subject types across treatments. Type 1 subjects are defined as those who choose not to draw a ball both when it is free and costly. Type 2 subjects are defined as those who choose to draw a ball when it is free but rather not draw a ball when it is costly. Type 3 subjects are defined as those who always choose to draw a ball. Type 0 subjects are inconsistent. They choose to draw a ball when it is costly but not when it is free. In practice, we do observe Type 2 subjects drawing some balls in the Costly treatment.

	Free	Costly	Forced
N	43	40	36
% Type 0	2%	0%	3%
% Type 1	14%	15%	17%
% Type 2	47%	43%	42%
% Type 3	37%	43%	39%

Table 5: Type distribution by treatment

Figure 6 presents aggregate average decisions by subject type, treatment and ball draw.[18] Note that this stratification lowers our sample sizes and results in a lower power of statistical tests. As a result we refer in this sub-section to the 10% level as our significance threshold. Average decisions of Type 3 subjects after a White draw are significantly lower in the Forced and Costly treatment than in the Free treatment (MWW: $p = 0.08$ and $p = 0.09$, respectively). No significant

[17] We use data for the 119 subjects who participated in sessions with the Identification block. Moreover, some subjects never drew a ball, and thus the number of subjects used is $N = 99$ for a Black draw and $N = 96$ for a White draw.

[18] Type 0 and Type 1 are not presented since our sample size for these subject types is very small. See Table 11 in the Appnedix for aggregate means of all subject types with standard errors.

differences are found for Type 3 subjects after a Black draw. Though the differences in Type 2 decision averages after a White draw between the Free and Costly treatments and the Forced treatment are graphically visible, they are not statistically significant. Type 2 differences in average decisions between the the Free and Costly treatments and the Forced treatment after a Black draw are borderline significant (MWW: $p = 0.11$ and $p = 0.12$, respectively).

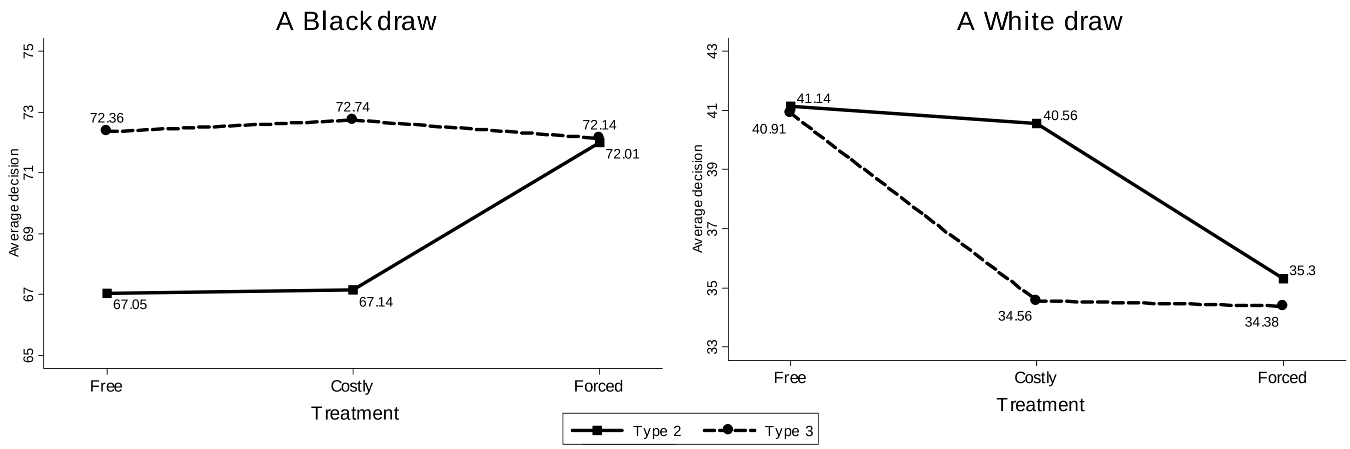


Figure 6: Average decision by subject type.

Focusing our attention on Type 2 and Type 3 subjects, and based on the non-parametric analysis, we find evidence of some difference in behavior in Costly. Type 2 individuals, when deciding to purchase information, act similarly in Free and Costly. Type 3, on the other hand, act differently (after a White draw) in Free and Costly. Still, in the Forced treatment both types make similar decisions. If selection had been the cause of the difference between Free and Costly, the average decision of Type 3 subjects would be similar. This is the case after a Black draw but not after a White draw. We now turn to parametric analysis to further our understanding of subjects' types on decision making. The following model is estimated:

$$\bar{d}_{i\phi} = \alpha + \beta_1 \cdot D_{i\phi}^{Costly} + \beta_2 \cdot D_{i\phi}^{Costly} \cdot D_{i\phi}^{Type\ 2} + \beta_3 \cdot D_{i\phi}^{Forced} + \beta_4 \cdot D_{i\phi}^{Type\ 1} + \beta_5 \cdot D_{i\phi}^{Type\ 2} + X_{i\phi}' \cdot \gamma + \epsilon_{i\phi} \quad (2)$$

The dependent variable $\bar{d}_{i\phi}$ and the subscripts i and ϕ are as defined in equation (1). D^{Costly} and D^{Forced} are treatment dummies and $D^{Type\ 1}$ and $D^{Type\ 2}$ are subject type dummies (X includes a gender dummy, $Female$, mathematical knowledge, $Math$, and risk aversion level, $Risk$). Table 6 presents OLS regression coefficients over decisions made after a Black draw and after a White draw for different independent variables. Columns (4) and (8) present results with the addition of an interaction term between Costly and Type 2 (CT2). The baseline is the Free treatment and Type 3 subjects. The results are consistent with the analysis done in Sub-section 3.2 for the effect the Costly and Forced treatments. Costly and Forced treatments shift decisions upwards after a

Black draw and downwards after a White one.[19]

	A Black draw				A White draw			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Costly	1.88	2.51	1.32	2.99	-6.05^{**}	-6.79^{***}	-6.44^{***}	-8.16^{***}
	(2.34)	(2.13)	(1.91)	(2.31)	(2.49)	(2.39)	(2.29)	(3.07)
CT2	–	–	–	−3.58	–	–	–	3.65
	–	–	–	(3.43)	–	–	–	(4.00)
Forced	3.92^{*}	4.67^{**}	3.46^{*}	3.47^{*}	-6.94^{***}	-7.42^{***}	-7.28^{***}	-7.32^{***}
	(2.17)	(2.27)	(2.09)	(2.09)	(2.54)	(2.49)	(2.48)	(2.51)
Type 1	–	–	-13.20^{***}	-12.90^{***}	–	–	9.68^{**}	9.45^{**}
	–	–	(3.62)	(3.64)	–	–	(4.54)	(4.55)
Type 2	–	–	-3.01^{*}	−1.92	–	–	1.18	0.13
	–	–	(1.80)	(2.19)	–	–	(2.15)	(2.60)
Constant	67.94	72.06	75.09	74.56	41.98	41.09	39.47	39.94
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
N	99	99	99	99	96	96	96	96

Table 6: Regression results by ball draw. Note: Type 0 subject decisions and outliers are not included in the analysis. Robust standard errors used. ***/**/* indicates significancelevel at the 1%/5%/10%.

The effect of the Forced treatment is stronger than the Costly treatment in both draws though the difference is not found to be significant (Wald test, $p = 0.33$). This is the opposite effect we would expect to find if subject selection in the Costly treatment played a large role. If the difference in average decision making between Free and Costly was driven by those subjects who choose to purchase costly information then any change we observe between these treatments must be larger than the difference between the Free and Forced treatments. The reason is that the average decision making in the Forced treatment is the weighted average of those subjects who choose to buy costly information and those who do not. It is thus the cost of information itself that directly affects subject behavior.

Result 3: *Differences in decision making are not driven by selection.*

Using Figure 6, we can attempt to explain why a stronger deviation is observed in Forced than in Costly. To do so, we focus our attention on the behavior of Type 2 subjects. Type 2 subjects' average decision after both a Black and a White draw are similar in the Free and Costly treatments and only shift in the Forced treatment. This is consistent with our type dichotomy if we consider Type 2 subjects' behavior in the Costly treatment as an exploratory one, investigating the benefits and costs of opting for a ball draw. Such 'testing of the waters' may be less affected by the cost of drawing a ball. Average decisions for Type 2 and Type 3 subjects should then be equal in Free and Forced while being different in Costly. This is indeed the case after a White draw but not

[19]See Table 12 and Table 13 in the Appendix for regression results including controls and for decisions made after No draw.

entirely after a Black draw.[20] The weak effect of Costly in Table 6 is thus the average effect of Type 2 and Type 3 subjects. This also justifies the significant effect of the Type 2 dummy seen in Table 6. Adding an interaction term between Type 2 subjects and the Costly treatment to the estimation of the model supplements the graphical explanation (see Table 6, Columns 4 and 8). The shift in decision making after both Costly and Forced becomes highly similar. Additionally, the Type 2 dummy coefficient after a Black draw is no longer significant. Table 6 also exhibits significant differences between Type 3 and Type 1 subjects' performance. Using a Wald test we also find significant differences between Type 2 and Type 1 subjects (p-values: 0.09 after a White draw and 0.01 after a Black draw).

Result 4: *Subjects who use freely available information (Type 2 and Type 3 subjects) perform similarly in the Free and Forced treatments and differently from subjects who do not (Type 1). Type 1 subjects substantially under-weigh new information.*

The experimental results discussed in Sub-section 3.2 and Sub-section 3.3 are closely related. We show that the variance in the cost of information is the driver of our result using both the Forced treatment and by using the Identification block data. We are thus able to consider the effect that subject heterogeneity, with respect to information purchasing decisions, has on our results. The change in behavior due to the cost of information is significant and systematic. Sunk cost effects are thus shown to have an effect on decision making in our experimental setting.

4 Pricing Information

In general, reliable and useful information is a lever towards better results. Our choice framework presents subjects with the opportunity of increasing expected gains through the incorporation of new information in a Bayesian fashion. In Sub-section 3.2 we observed that subjects put relatively more weight on information they had to pay for, but this does not always lead to better decisions. However, even if information always pushed subjects closer to the optimum, this tells us little on the efficiency of using information. In other words, the gains realized from using information (because of better decisions via reduced uncertainty) might not compensate the cost paid for it.[21] In this section we investigate this implicit trade-off by means of a small exercise that tells us how information should have been priced (or subsidized) in order for it to be profitable for subjects. Note that in this section we move our focus away from the effect of sunk costs to the efficiency gains obtained from using information.

To answer this question we compute the cost levels that would make subjects indifferent between paying for information and having no information. We use data from the Decision block in this

[20] After a Black draw Type 3 subjects average decision is significantly higher than Type 2 subjects in the Free treatment (MWW: $p = 0.07$).

[21] Recall that we priced information at roughly 60% of the expected gain observed in the Bayesian benchmark.

analysis, and only from the Free and Forced treatments (the selection present in Costly complicates data interpretation, as many subjects chose to never observe observation - cf. Sub-section 3.1). We want to calculate the individual cost levels that make the following equality hold:

$$\sum_{\phi \in \{Black, White\}} \Pr(\phi)\, V\left(\bar{d}_{i\phi}\right) - c_i = V\left(\bar{d}_{i,No\ Info}\right)$$

where $V()$ is the payoff that would result by implementing the average decision(s) of each subject in the respective information condition, $\bar{d}_{i,\phi}$ (defined in equation 1). Figure 7 presents a plot of the implied cost levels, c_i, in ascending order. To be more precise, c_i is the cost level which would make subjects indifferent between facing one decision with paid information and one decision without information. This value is conditional on the average decisions of each subject in each information condition. We observe that the great majority of subjects should be willing to pay for information ($c_i \geq 0$): only 10% would have to be subsidized ($c_i < 0$). Moreover, 60% of subjects have an implied cost level above 0.3, which means that information was priced in a beneficial way for the majority of subjects. Based on their implied maximum willingness to pay for information, did subjects select themselves correctly? Restricting the sample to observations with Identification block data, an ordered logit regression of type on implied cost levels shows that there is a positive and significant relationship between the two variables (see Table 7). This means that a higher implied cost level tends to be associated with a higher subject type.

A noteworthy aspect is the fact that Forced did not lead to a better overall use of information, as the distribution of c_i in this treatment follows the one in Free quite closely. The reason is that, as mentioned, subjects in Forced got very close to the optimum in case of a Black ball but overshoot in case of a White ball (see Figure 4).

	Type
Implied Cost Level	2.06**
	(0.95)
Controls	Yes
Observations	69
Log-pseudolikelihood	−64.78

Table 7: Regression results: types and implied cost levels. Note: Ordered logit regression. Robust standard errors in parentheses. ** indicates significance at the 5% level.

Despite the numerous caveats inherent to this exercise, the main message is that information pricing is far from trivial from a policy perspective. On the one hand, it can provide (the right) incentives if it leads to a better incorporation of information in decision making. On the other hand, individuals might end up worse off if the price paid for information cancels the benefits derived from having information available. However, the evidence suggests that this trade-off seems to be solved to some extent by allowing individuals to decide whether they want to acquire information or not. All in all, this is a point where further research is needed.

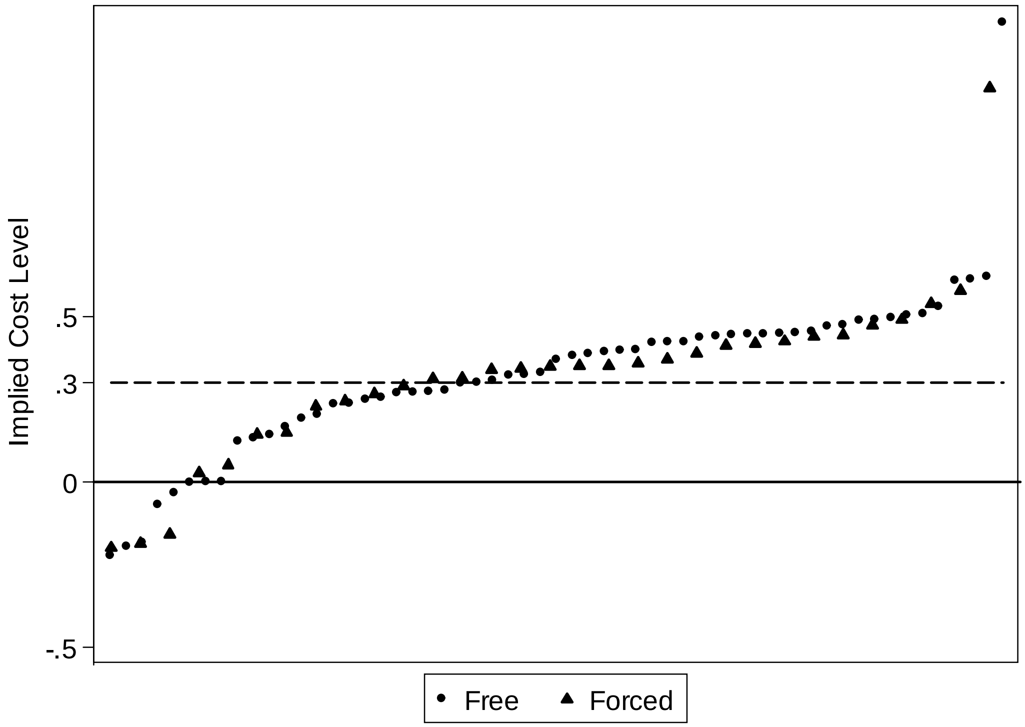


Figure 7: Implied cost levels. Note: $N_{Free}= 57$and $N_{Forced}= 31$as some subjects did not see a Black or a White balland. The observations of each treatment are equally spaced over the axis. The dashed line (0.3) is the price of information in the Decision block.

5 Conclusion and discussion

This paper sets to explore how individuals' 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 updating of beliefs, 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 at no cost and after paying for new information. Consequentially, sunk cost may have an affect on individuals' deviation from Bayesian updating. Biases in the updating of beliefs may thus be exacerbated or alleviated by cost. To examine these issues we use a laboratory experiment which enables us to control for the selection problem. That is, we control for the possibility that any effect we detect in the data is engendered by the difference in behavior between subjects who have high value of information and those with low value of information, rather than by the cost of information itself. We do so using two independent procedures. First, we implement a treatment in which we force subjects to pay for information regardless of their will to use it. Second, we identify subjects' demand for costly information regardless of the treatment. We can then compare those subjects who choose to buy information when its costly with those who would counterfactually do so. We find significant sunk costs effect on decision making by the subjects participating in the experiment. Subject 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 beliefs towards the extremes. Decision making can be closer to or further from the correct Bayesian updating, conditional on decision making 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.

Since placing a cost on information can be beneficial or harmful, this paper suggests that policy makers should perhaps consider the implication of costly information. It can be favorable as long as individuals under-update new information when it is given freely. The down side of making information costly is that a trade-off exists between better individual updating and lower demand for information. Another possibility is for a policy maker to force a cost from a mandatory fee or tax on information. This fee must be directly related to the information such that an individual does not internalize the fee in advance and ignore its relevance for information. Our results may also be of importance for research on the value of information. Economists sometimes vary the cost of information across treatments in order to elicit subjects' value of information. Our paper shows that this costly information can affect decision making after information is given.

As we now have evidence on the existence of cost effects on the use of information, further examination of possible interaction effects, as well as existence in other environments, may be of interest. A natural extension of this paper is to add treatments with differing information costs. For example, if a low cost is found to induce similar changes in behavior as the ones found in this paper then using costly information may be useful also in cases where demand for information is sensitive to price. Another interesting extension would be to examine the effect of sunk cost on information in a setting involving strategic interactions. A design that would allow an examination of the exact channel through which the sunk cost effect operates would also be of high interest.

References

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. *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.

Camerer, C. (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. (1987). Do biases in probability judgment matter in market? experimental evidence. *American Economic Review 77*(5), 981.

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 and Organization 39*(1), 59 – 82.

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

Cohen, J. and P. Dupas (2010). Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. *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), pp. 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), pp. 631–645.

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

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

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.

Koehler, J. J. (1996, June). 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.

Offerman, T., J. Sonnemans, G. Van De Kuilen, and P. Wakker (2009). A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes. *The Review of Economic Studies 76*(4), 1461–1489.

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.

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), pp. 474–494.

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

6 Appendix

6.A Details on the Choice Framework and Type Classification

6.A.1 Choice Framework

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

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

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

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

In our experiment $x \in [0, 100], l = 20$ and $r = 80$. As explained in Sub-section 2.1 there are three possible information conditions, $\phi \in \{Black, White, No\ Info\}$, which induce different distributions of the signal. We define a "$Draw$" as a "$Black$" or a "$White$" ball draw. Table 8 presents the two parameterizations that were implemented: A was used in the Decision block and B was used in the Identification block.

	α	β	γ	$\Pr(L)$	$\Pr(Black\vert L)$	$\Pr(Black\vert R)$	Cost	Exch. Rate
A	6	0.009	1.7	0.4	0.2	0.6	0.3	0.75
B	5.7	0.00925	1.7	0.7	0.8	0.4	0.25	0.75

Table 8: 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 $E[F(x^*, \sigma)|Draw]$, is provided for both parameterizations in Table 9. Note that the scenarios induce similar expected values in all information

	$\Pr(L\vert\phi)$		$x^*\vert\phi$			$E[F(x^*,\sigma)\vert\phi]$			$\Pr(Black)$	$E[F(x^*,\sigma)\vert Draw]$
	Black	White	No Draw	Black	White	No Draw	Black	White		
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 9: Values for Parameterizations A and B.

conditions. This makes the incentive to optimize and acquire information similar across parameterizations. Notwithstanding, the scenarios look sufficiently different from each other such that the rules employed in one are not easily translated to the other. Note that the prior probability changes while the urn composition remains unchanged (there is a mere relabeling of colors and states).

In reality, individuals will have their own estimates of the (updated) probability p, $p_i = p_i(\phi)$. This will lead to individual decisions $x_i^* = x(p_i)$. Therefore, we define the utility value derived from the observed payoff as $U_i = u(F(x_i^*, \sigma))$, where $u()$ is a general utility function. Given that u_i only depends on the primitives ϕ and σ, we write $U_i(\phi, \sigma)$.

6.A.2 Type Classification

We define our types according to subjects' willingness to buy information. A subject buys information if:

$$V_i(Draw) - C_{1,i} - C_2(\theta) \geq V_i(No\ Info) \tag{A.1}$$

where

$$\begin{aligned}
V_i(Draw) &= E(U_i(Draw, \sigma)) \\
&= p_i [\Pr(White|L) U_i(White, L) + \Pr(Black|L) U_i(Black, L)] \\
&+ (1 - p_i) [\Pr(White|R) U_i(White, R) + \Pr(Black|R) U_i(Black, R)] \\
V_i(No\ Info) &= E(U_i(No\ Info, \sigma)) \\
&= p_i U_i(No\ Info, L) + (1 - p_i) U_i(No\ Info, R)
\end{aligned}$$

We can re-write equation A.1 as:

$$\begin{aligned}
&p_i [\Pr(White|L) U_i(White, L) + \Pr(Black|L) U_i(Black, L) - U_i(No\ Info, L)] \\
&+ (1 - p_i) [\Pr(White|R) U_i(White, R) + \Pr(Black|R) U_i(Black, R) - U_i(No\ Info, R)] \\
&\geq C_{1,i} + C_2(\theta)
\end{aligned}$$

which tells us that a subject acquires information if her estimate of the expected gain when information is available is sufficiently higher that her expected gain with no information. Our specification makes use of a couple of assumptions. First, both types of cost are separable from benefits in the utility function and they are linearly additive. Second, cognitive costs for No Draw are normalized in such a way that we can write $C_{1,i} = C_{1,i}|Draw - C_{1,i}|NoDraw$.

6.B Experiment Instructions

Below we provide an abridged transcript of the instructions. Square parentheses indicate changes in the sessions with a different Identification block.

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 2.1 was shown here, see a graphical representation of the urns in Appendix 6.D)

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 2.1 was shown here, see a graphical representation of the urns in Appendix 6.D)

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.]

6.C Additional results

Table 10 presents the p-values of the Kolmogorov-Smirnov test for equality of distribution. It is applied to the differences in individual average decision distribution across treatment.

Table 11 presents average aggregate decision by subjects types and ball draw.

Table 7 is identical to Table 6, with the control variables explicitly shown. Math has a significant effect after both draws. A larger number of math courses leads to a shift towards less extreme

	A Black draw	A White draw	No draw
Free vs. Costly	0.05	0.03	0.57
Free vs. Forced	0.136	0.01	0.15
Costly vs. Forced	0.87	0.34	0.54

Table 10: Kolmogorov-Smirnov tests of decision distributions

	Black draw			White draw			No draw		
	Type 1	Type 2	Type 3	Type 1	Type 2	Type 3	Type 1	Type 2	Type 3
Free	57.15	67.05	72.35	52.94	41.14	40.91	56.91	54.31	56.71
	(5.63)	(2.17)	(2.00)	(7.88)	(2.26)	(2.83)	(3.1)	(1.13)	(1.53)
Costly	80.5	67.14	72.74	39	40.55	34.56	52.31	55.38	58.06
	(19.5)	(2.60)	(2.19)	(−)	(4.02)	(2.62)	(1.64)	(1.91)	(1.15)
Forced	65.5	72.01	72.14	37.67	35.30	34.28	63.33	57.13	52.07
	(−)	(2.09)	(2.20)	(3.17)	(3.16)	(2.24)	(4.04)	(1.35)	(2.09)

Table 11: Decision means by types. Standard errors in brackets. Empty brackets indicate only one observation.

decision making. Gender has a significant effect only after a Black draw. Female subjects also tend towards less extreme decisions. Risk does have a significant effect. It is worth mentioning that our measure of risk is truncated at risk neutrality ($Risk \in [0, 3]$ where $Risk = 3$ is risk neutrality). Our measure can thus not detect risk-loving behavior.

Table 13 expands the analysis shown in Table 7 to No draw data. No variable is significant but the Gender control. female subjects tend to make decisions closer to 50 than the male subjects.

	A Black draw				A White draw			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Costly	1.88	2.51	1.32	2.99	-6.05^{**}	-6.79^{***}	-6.44^{***}	-8.16^{***}
	(2.34)	(2.13)	(1.91)	(2.31)	(2.49)	(2.39)	(2.29)	(3.07)
CT2	–	–	–	-3.58	–	–	–	3.65
	–	–	–	(3.43)	–	–	–	(4.00)
Forced	3.92^{*}	4.67^{**}	3.46^{*}	3.47^{*}	-6.94^{***}	-7.42^{***}	-7.28^{***}	-7.32^{***}
	(2.17)	(2.27)	(2.09)	(2.09)	(2.54)	(2.49)	(2.48)	(2.51)
Type 1	–	–	-13.20^{***}	-12.90^{***}	–	–	9.68^{**}	9.45^{**}
	–	–	(3.62)	(3.64)	–	–	(4.54)	(4.55)
Type 2	–	–	-3.01^{*}	-1.92	–	–	1.18	0.13
	–	–	(1.80)	(2.19)	–	–	(2.15)	(2.60)
Math	–	-1.14^{***}	-1.31^{***}	-1.34^{***}	–	1.44^{***}	1.58^{***}	1.61^{***}
	–	(0.37)	(0.36)	(0.36)	–	(0.45)	(0.45)	(0.45)
Female	–	-4.58^{***}	-3.35^{***}	-3.37^{*}	–	-0.06	-0.50	-0.47
	–	(1.84)	(1.71)	(1.74)	–	(1.98)	(2.03)	(2.04)
Risk	–	0.23	0.22	0.24	–	-1.21	-1.16	-1.16
	–	(0.97)	(0.95)	(0.95)	–	(1.25)	(1.27)	(1.26)
Constant	67.94	72.06	75.09	74.56	41.98	41.09	39.47	39.94
N	99	99	99	99	96	96	96	96

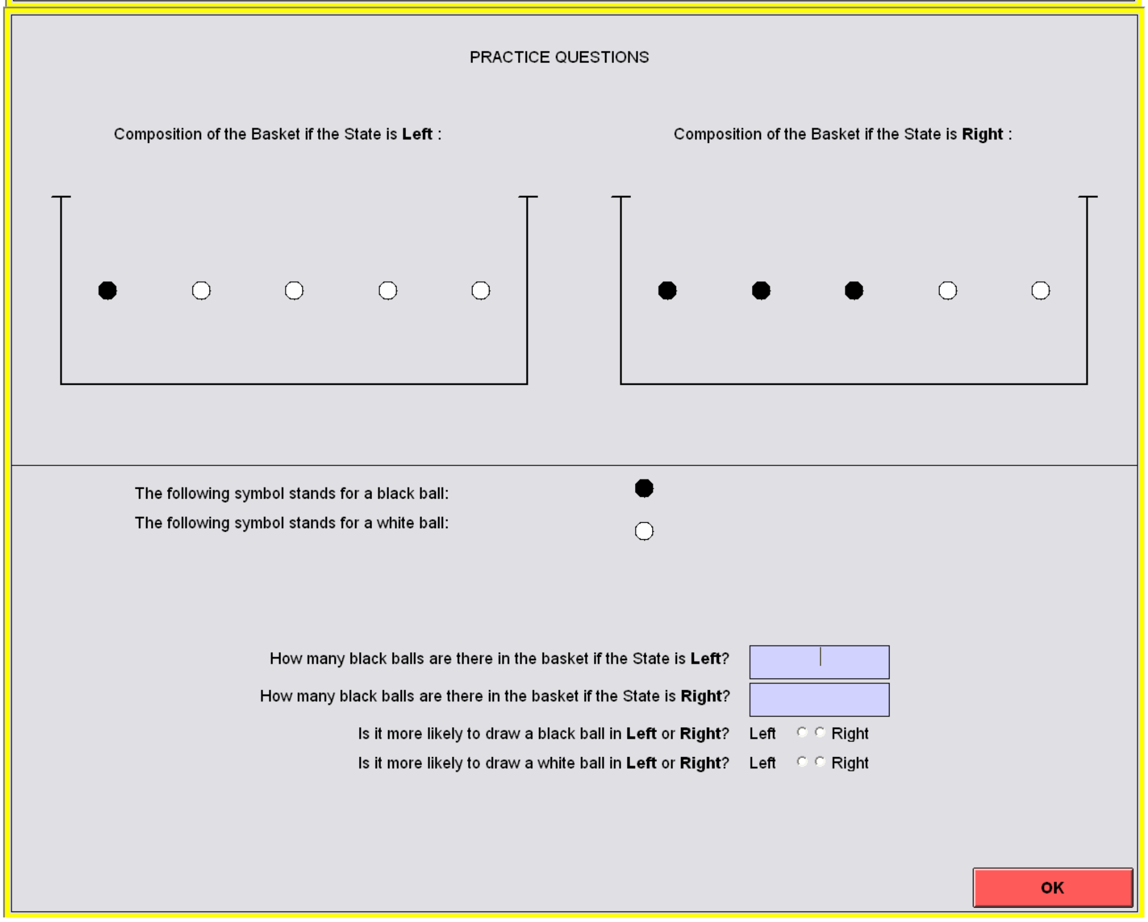
Table 12: Regression results by infomation. Note: Type 0 subject decisions and outliers are not included in the analysis. Robust standard errors used. ***/**/* indicates significancelevel at the 1%/5%/10%..

	No draw		
	(1)	(2)	(3)
Costly	0.46	0.89	0.91
	(1.36)	(1.24)	(1.30)
Forced	0.57	1.00	0.99
	(1.64)	(1.66)	(1.60)
Type 1	–	–	2.86
	–	–	(2.12)
Type 2	–	–	0.78
	–	–	(1.28)
Math	–	-0.32	-0.29
	–	(0.27)	(0.28)
Female	–	-2.68^{**}	-2.97^{**}
	–	(1.33)	(1.36)
Risk	–	0.78	0.92
	–	(0.72)	(0.71)
Constant	55.60	57.56	57.03
N	117	117	117

Table 13: Regression results by infomation. Note: Type 0 subject decisions are not included in the analysis. Robust standard errors used. ***/**/* indicates significancelevel at the 1%/5%/10%..

Appendix 6.D – Snapshot of the Experiment

Below we reproduce snapshots of the practice questions:



Below we reproduce a snapshot of the experiment (Free treatment):