Rational inattention and timing of information provision [1]

Diego Aycinena
Department of Economics,
Universidad del Rosario and
Economic Science Institute, Chapman University
diego.aycinena@urosario.edu.co

Alexander Elbittar
Department of Economics
CIDE
alexander.elbittar@cide.edu

Andrei Gomberg
Department of Economics
ITAM
gomberg@itam.mx

Lucas Rentschler
Department of Economics and Finance
Utah State University
lucas.rentschler@usu.edu

June 17, 2020

[1]We gratefully acknowledge participants at the 2018 North American Economics Science Association Conference in Antigua, 2018 Meetings of the Society for Social Choice and Welfare, 2018 Utah Experimental Economics Conference, the 2019 Bogotá Experimental Economics Conference, seminars at George Mason University, Florida International University and CIDE. Simón Caicedo provided wonderful research assistance, and Mario Sandari Gomez and Megan Luetje provided invaluable research support with experimental sessions. Gomberg would like to acknowledge the financial support of Asociación Mexicana de Cultura. Aycinena would like to acknowledge the generous financial support from the Economic Science Institute at Chapman University.

Abstract

We consider the issue of how timing of provision of additional information affects information-acquisition incentives. In environments with costly attention, a sufficiently confident agent may choose to act based on the prior, without incurring those costs. However, a promise of additional information in the future may be used to encourage additional attentional effort. This may be viewed as a novel empirical implication of rational inattention. In a lab experiment designed to test this theoretical prediction, we show that promise of future "free" information induces subjects to acquire information which they would not be acquiring without such a promise.

JEL Classifications: C72, C90, D44, D80.
Keywords: information acquisition, rational ignorance, experiments.

1 Introduction

Providing "cheap" or "salient" information has emerged as a major theme in the literature on consumer choice (for a recent review, see Bernheim and Taubinsky, 2018). Bollinger et al. (2011) studied the impact of providing caloric information on purchases by Starbucks clients; Chaloupka et al. (2015) considered the impact of warning label regulation on smoking; Allcott and Taubinsky (2015) showed how providing comparative information on products affects consumer switch between light bulb technologies, and Bhargava et al. (2017) explored the impact of clarity improvements in presentation of health insurance plans. Information provision has also been shown to affect political choices. Thus, León (2017) studied the impact of information on monetary penalties for abstention on turnout in Peruvian municipal elections and Larreguy et al. (2020) produced evidence of impact of dissemination of information on government performance on voting outcomes in Mexico. The role of information provision in "nudging" behavior appears by now well-established. Yet, little has been done to explore the interaction between information provided and individuals' own information acquisition effort. In particular, we are not aware of any studies that focused on the role played by timing of information provision. This is the gap that we address in this paper.

It may, at first glance, appear that the promise of free information in the future might merely crowd out independent information acquisition by an agent. However, it turns out that a promise of additional information in the future may, in fact be used to incentivize individual information acquisition effort. The reason for this could be found in the well-known observation by Radner and Stiglitz (1984) that marginal value of small amounts of information is usually zero. Intuitively, this arises from the fact that weak signals do not affect optimal behavior and are payoff-irrelevant ex post, no matter their realization. Chade and Schlee (2002) have shown that the resultant (Radner-Stiglitz) non-concavity in the value of information is an extremely robust feature of costly information acquisition environments. It implies that individual demand for information would normally exhibit discontinuities at zero: there is a minimum scale at which information should be acquired, and people may want to stop paying attention and act based on their prior beliefs, unless reaching this threshold is sufficiently cheap.[1]

It turns out that a promise of additional information to be delivered after an individual completes her own information acquisition may be used to "smooth out" this non-concavity. Indeed, if an agent knows that additional information will be provided, but she has to decide on her own effort before observing whatever may be told to her later, anything she learns on her own would be valuable with positive probability. As long as future help is expected, marginal value of small amounts of information is no longer zero. Rather than substituting for individual effort, promised free information may complement it. Hence, such promise may indeed stimulate acquiring small amounts of information, which otherwise should never be acquired.[2]

This straightforward and intuitive point has not, to the best of our knowledge, been made or tested in the literature. It is, however, a common implication of costly information acquisition environments. In this respect, our work is related to the recent theoretical and experimental work on rational inattention (Caplin and Martin, 2015; Caplin and Dean, 2015; Caplin and Martin, 2018; Dean and Neligh, 2017). In common with these studies, we explore consequences of agents rationally deciding how much attention they would

[1] There is, of course, an obvious situation when this is not the case, but it is non-generic: if without additional information the agent is indifferent between actions, arbitrarily small amounts of information could be sufficient to break this indifference. This, however, is readily seen to be a rare knife-edge case.

[2] Substitution intuition is not completely wrong: if the agent is indifferent between actions, arbitrarily small amounts of information become valuable. In this rare knife-edge case she may hope to free ride on any promised future help, thus, crowding out her effort. As this argument relies on initial exact indifference, it would not be something one could expect to easily observe in real life. But this this allows us to introduce a subtle test in the lab in order to provide a further test for the theory.

like to pay in a costly information environment. Promising additional information in the future allows us to directly vary the set of available informational strategies. In this sense, the interaction between individual information-acquisition effort and timing of free information provision we explore here may be seen as a novel implication of rational inattention.[3]

Though a promise of free information may create incentives for costly information acquisition—that would not exist otherwise, it would not improve the quality of individual decision, nor the welfare of the decision-maker. This result follows straightforwardly from Blackwell (1962). In some situations, however, we may expect greater individual learning to have spillovers on others. In fact, this reasoning may be behind the typical prohibition for jurors to talk about the evidence they listen to during the course of a trial. Jury discussion is supposed to occur only after they listen to every bit of evidence presented. In Hannaford et al. (2000), for example, the authors show intriguing evidence of increased drop out by jurors in a field experiment in which these instructions were relaxed, which lends support to the hypothesis we have just laid out.

Our experimental design follows the approach of recent laboratory studies on costly information acquisition (Bhattacharya et al., 2017; Elbittar et al., 2016; Grosser and Seebauer, 2018).[4] It allows us to explore the predictions of the model using both between- and within- subjects data. Our results are consistent with the predictions of the model: offering future information encourages greater investments in information acquisition. Furthermore, when we explore the rare knife-edge case of indifference, we find, as predicted, that information acquisition reverses with the promise of future information.

In section 2 we provide a theoretical model of attentional response to variation in timing of free information provision. Section 3 introduces the experimental design and section 4 describes our data and empirical analysis. To encourage you to read along, we promise to report the results of the experiment and conclusions in the last two sections.

2 Theory

Following a simplified version of the Radner and Stiglitz (1984) set-up[5], let Ω be a finite set of states of the world, S be a finite set of signals, and $\Theta = [0, 1]$ be the set of available *signal structures*. For each $\omega\ 2\ \Omega, s\ 2\ S$ and $\theta\ 2\ \Theta$ let $p_{\omega,s}(\theta)$ denote the conditional probability of signal s given the state ω. Let $p_{\omega,s}(0) = p_s$ be independent of ω (corresponding to a completely uninformative signal) and we shall assume that $p_{\omega,s}(\theta)$ is continuous and differentiable at $\theta = 0$ for each $\omega\ 2\ \Omega$ and $s\ 2\ \Theta$, implying that the informativeness of the signal structure changes smoothly around $\theta = 0$. Let A denote a (compact) action space and suppose the agent has some state-contingent continuous Bernoulli utility $u_\omega(a)$ and let β_ω denote the prior probability of state ω. Then, upon observing signal $s\ 2\ S$, she would maximize expected utility

[3]A somewhat similar effect has been noted by Caplin and Martin (2018) who explored, from a rational inattention standpoint, how varying default options presented to subjects may nudge them to either acquire information or to "drop out".

[4]These studies in juries establish a framework in which potential informational spillovers from increased study generated through delayed communication between jury members may be explored. It is of interest that, in the last two studies referenced, the subjects consistently acquired less information than predicted for the experimental setting. Elbittar et al. (2016) propose that this may arise from biased priors, which effectively lead to the drop out from information acquisition of the sort we study in this paper. Offer of future information, which in this setting may be interpreted as arising from jury deliberation, would be expected to help resolve this problem.

[5]While Radner and Stiglitz have established the result for finite state and signal space - an environment we also consider in this paper, (Chade and Schlee, 2002) have since shown that it remains pervasive in the infinite settings as well.

and take an action

$$a(s,\theta) \in \arg\max_{a \in A} \sum_{\omega} \beta_\omega p_{\omega,s}(\theta)\, u_s(a)$$

We can then define the *ex ante* value of the signal structure $\theta \in \Theta$ to be

$$v(\theta) = \sum_{\omega,s} \beta_\omega p_{\omega,s}(\theta)\, u_s(a(s,\theta))$$

Then, assuming $a(.,\theta)$ is a continuous function in θ for each $s \in S$ and $p_{\omega,s}(\theta)$ is continuous and differentiable at $\theta = 0$ for each $\omega \in \Omega$ and $s \in \Theta$, the main theorem of (Radner and Stiglitz, 1984) states that $\limsup_{\theta \to 0} \frac{v(\theta) - v(0)}{\theta} = 0$: the marginal value of information, if it is defined, $v'(0)$ is zero. Hence, if there is any cost of acquiring information (that is moving from $\theta = 0$ to some $\theta' > 0$), then our agent will choose not to acquire small amounts of it: there is a minimum amount of information $\bar{\theta} > 0$ that may be acquired by an agent maximizing her *ex ante* expected utility.

Suppose, however, that the agent knows that, after she is done with her information acquisition, she will observe additional information, moving her to $\theta > 0$ *at zero cost*. The nature of her information-acquisition decision is now changed: a small additional amount of information acquired will have marginal value $v'(\theta)$, which may be positive. Hence, it is now possible that even smallest amounts of information would be optimal for the agent to acquire. It is this insight that we experimentally exploit in this paper.

In our experimental design we consider a further simplified setting with two states and two actions. Specifically, the two possible states of the world are $\Omega = \{R, B\}$. The agent assigns a prior probability $P(R) = \beta \in (0,1)$ to $\omega = R$ and chooses $a \in \{r, b\}$.

We assume that the agent is indifferent between the two states, as long as her decision is correct ($U(R,r) = U(B,b) = 0$), while her attitude towards the two possible errors may differ, so that $U(B,r) = -q$ and $U(R,b) = -(1-q)$ so that $q \in (0,1)$ may be interpreted as the degree of certainty necessary for the agent to choose $a = r$. Indeed, with these parameters the agent would be willing to choose $a = r$ if and only if $P(R) \geq q$.

Each signal observed by the agent is binary, $S_i \in \left\{\hat{r}, \hat{b}\right\}$ and is correlated with the true state of the world, so that $P(S_i = \hat{r} | \omega = R) = P\left(S_i = \hat{b} | \omega = B\right) = p \in \left(\frac{1}{2}, 1\right)$. Conditional on the true state, signals are *i.i.d.* Before choosing a, the agent decides how many signals she would like to purchase with a constant per signal cost $c > 0$. In addition a fixed number of free signals will be observed before the agent chooses v, and the agent knows this. We concentrate on the interaction between the timing at which the free signals are observed by the agent and her decision to purchase additional signals.

It is useful to consider the special case, in which before any information arrives the agent's prior belief $\beta = q$, so that she is indifferent between the two states. In this case standard Bayesian updating implies that she should choose $a = r$ whenever $m \equiv \#\{i : S_i = \hat{r}\}$ is bigger than $k \equiv \#\left\{i : S_i = \hat{b}\right\}$ and should prefer to declare $a = b$ whenever $k > m$ (the agent is indifferent whenever $k = m$). In this case, therefore, the standard Bayesian updating coincides with the simple "count-the-signals" rule of thumb, greatly simplifying the decision the agent faces. Notably, however, this is also the case in which even minute amount of information would break the indifference, so marginal value of information at zero is, in fact, positive (this is possible since in this - and only this - case action is not continuous in information). This, in turn, makes possible a reversal of the expected impact of free information delay: there is no longer any non-concavity for the promise of future information to smooth, but the agent may be tempted to free ride on this promise. The particular situation, of course, is non-generic, but we shall be making an extensive use of this environment in our design.

As discussed above, except in the knife-edge indifference case, we may be able to induce greater information acquisition by promising that more information will be given after the agent purchases information; however, the additional learning will not improve the expected quality of the verdict. To see this, suppose a total of n signals, $S_1^f, ...S_n^f$ may be provided to the agent for free. If we choose to provide them at the beginning, so that their realizations may be observed before the agent decides how many signals to purchase at cost, she would be able to condition her decision on the observed realizations. Defining the difference between the number of signals that indicate R and B as $X^f = \#\left\{i : S_i^f = \hat{r}\right\} - \#\left\{i : S_i^f = \hat{b}\right\}$, we observe that, in order for the additional information to have any potential impact on the agent's choice of v, she would have to buy more signals than the the observed realization of X^f. Indeed, suppose the number of costly signals she purchases, $s \leq |X^f|$. Then, even if all of the signal realizations are identical, the sign of $m - k$ is be equal to the sign of X^f, implying the same choice of v. This implies that the expected value of purchasing $s \leq |X^f|$ signals is equal to zero: she should either buy a lot of information, or none –the Radner and Stiglitz (1984) in action. Of course, even a single signal would be valuable if the free signals "tie", $X^f = 0$. This tie can only occur if the number of these signals n is even, in which case, irrespective of the true state of the world it would happen with probability $P\left(X^f = 0\right) = \frac{n!}{\left(\frac{n}{2}!\right)^2} p^{\frac{n}{2}} \left(1 - p\right)^{\frac{n}{2}} > 0$.

If, we instead promise to show the agent the same signals after she decides how many signals to purchase, this will be her prior probability that $X^f = 0$. Hence, with positive probability even a single signal would turn out to be valuable.

3 Experimental Design

Our experimental task follows a standard information acquisition environment (Elbittar et al., 2016; Guarnaschelli et al., 2018; Battaglini et al., 2010). Participants earn money if they correctly guess the true binary state of the world, $\Omega = \{R, B\}$, framed in the context of guessing the color of a jar, that is either Red or Blue.[6] Subjects received four free balls drawn from the jar with binary signals, S_i, corresponding to the two possible states, and could acquire additional signals at a cost. Our treatment variable is the timing of free signals relative to the acquisition decision.

In each period, participants know that the true color of a jar is equally likely to be Red or Blue. Jars contain 60 balls that correspond to the jar's color and 40 balls of the other color. Before guessing the color of the jar, the participant will see at no cost the color of 4 balls drawn from the jar independently and with replacement. In addition to those free balls, they have the opportunity to purchase up to 5 additional balls, at a known cost of c each. This cost is drawn in each period from a uniform distribution between 0 and 100.

We label our treatments according to the number of free balls to be observed *after* the information acquisition decision: *0B* (zero balls), *2B* (two balls), *4B* (four balls). In *0B*, all 4 free balls are shown *before* the decision to acquire information, so participants know that they will not observe any (0) free balls after their decision. In *2B*, 2 free balls are shown *before* and the other 2 free balls are shown shown *after* the decision to acquire additional information. Finally, in *4B*, all 4 free balls are shown *after* the decision to acquire

[6] In addition to the main task, there were three additional tasks that followed: a risk elicitation task using a multiple price list, and two additional jar-guessing tasks (for 12 periods each). Participants knew that one of the four tasks was to be randomly selected at the end, and their choice in a single (randomly selected) period for the task was to determine their earnings from the experiment. The experimental earnings was the sum of the randomly selected choice plus the show up fee and a starting balance of E$120.

information. The exogenous variation in the realization of free signals prior to information acquisition is a feature of our design which gives rise to different informational situations with different predictions.

Prior to guessing the color of the jar and after the information acquisition decision, participants see the color of all balls: the 4 free balls plus the number of balls acquired. Participants who guess correctly earn E$1,000 minus the cost of purchased balls; otherwise, they earned E$300 minus cost of purchased balls. They do this for 24 periods, all under the same treatment (between-subjects design). In the first twelve periods they do not obtain feedback from their guess; in other words, they do not learn the true state. Under the last twelve periods, they do observe feedback from their guess.[7]

We deliberately chose to induce payoffs and initial priors so that without any information subjects were exactly indifferent between the jars. As discussed in the previous section, this made the jar choice particularly simple: subjects needed merely to choose the jar corresponding to the color of the majority of the observed balls. As we demonstrate in the next section, the subjects used overwhelmingly that heuristic. This allowed us to concentrate on the information acquisition decision, which was our primary interest.

Table 1 presents the predicted marginal value of each additional signal across every informational situation in which a risk neutral participant may find herself. As previously mentioned, a feature of our experimental design is that the realized structure of free signals observed before the decision of information acquisition in each treatment exogenously induces different updated priors. These different informational situations (updated priors) lead to different predictions, as illustrated in the Table.

Due to the non-concavity in the value of information, the main prediction of the model is that, promising future -as in *2B* relative to 0B- information should induce greater information acquisition when the updated prior is asymmetric (i.e., $\beta_\omega = 9/13$). The opposite is true in the rare knife-edge case where the updated prior leads to indifference ($\beta_\omega = 1/2$). This is reflected in the marginal value of additional signals across the different treatments and induced updated priors. The intuition is the following: Assume that in 0B, where a participant observes four free balls before she decides how many to acquire at cost, she is shown at least 3 balls of the same color ($\beta_\omega = 9/13$). Acquiring two balls or less does not provide evidence sufficient to change beliefs. Thus, she would have to acquire at least 3 balls at cost for this additional information to have any value. Given the cost of this investment, such a participant is likely to "drop out" by purchasing no costly information at all. Contrast this with *2B* where a participant observed both balls of the same color ($\beta_\omega = 9/13$). In this case, purchasing a single additional ball would make sense as the future information (2 additional balls) can smooth out the non-concavity. Thus, whether she is in treatment *0B* and observed 3 Red balls or in treatment *2B* and observed 2 Red balls, a Bayesian subject should have identical prior about the likelihood of a Red jar at the time of the information acquisition decision. However, purchasing a single additional ball would only make sense in the latter case. Hence, in this setting ($\beta = 9/13$), we should observe fewer drop out decisions in *2B* as compared to *0B*.

An additional advantage of this set-up is the effective oversampling of the otherwise rare knife-edge case ($\beta_\omega = 1/2$). In this rare case even small amounts of information matter and, hence, the effect of the timing of free information provision is reversed.

4 Data and Empirical Analysis

We conducted two waves of sessions. In the first wave, 72 undergraduates (mainly) from Universidad Francisco Marroquín participated in the experiment. Within each session, a third of the subjects were randomly

[7] Instructions are available on Appendix D. Appendix C provides a description of the experimental protocols.

assigned to one of the three treatments; that is, we implemented all treatments within same session (to different subjects). We used the same set of (pre-defined random) draws across treatments.[8] In the second wave, 84 undergraduates from Chapman University took part. In each of these session we used $i.i.d.$ draws for each individual decision. All subjects within a session participated in the same treatment for the first 24 periods. In addition, for tasks 3 and 4, we repeat the main jar-guessing task for the other 2 treatments.[9] Thus, in these sessions we have between-subject variation (in the first 24-period) task that is comparable with the first wave, and within-subject variation from tasks 1, 3 and 4.[10]

We exploit both the between- and within- subject variation in our data. We pool together data from the main task in the 24 periods across the two waves for our between-subjects (BSs) data. In addition, we take from the second wave the last 12 periods of the main task and the 12 periods of each of tasks 3 and 4. This is our within-subjects (WSs) data. We report data from 156 subjects who took part in 12 sessions: six 12-subject sessions from the first wave of data collection, and six 14-subject sessions for the second wave.

Given our design, there are some ancillary predictions that we can rely on to check whether participants understood the environment and behave consistently with a "rational" choice model –beyond the precise main predictions of the model. First, Bayesian subjects should make their final decision by simply counting the number of balls of each color they observe in total, and declaring the majority color as the color of the jar (they would be indifferent in case of a tie). They do so 94.4% (96.5%) of the time in the BSs (WSs) data. Second subjects react to information acquisition prices.[11] Finally, as Table 1 illustrates, an idiosyncratic feature of our experimental design is that purchasing a positive even number of balls is dominated. Since the last of these would never make the subject strictly prefer to change her decision, it would never have a positive value in expectation. We see that conditional on acquiring information, subjects tend to acquire odd number of signals 73.2% of the time (77.6% during the second half).[12]

We test our main hypotheses using both non-parametric tests and reduced form regressions on treatment effects. For our non parametric tests we take the average value of the variable of interest for each individual across all 24 periods (and use that to make comparisons across treatments (BSs). For the WSs data, we compare the average of value of the variable of interest for each individual across all periods of each treatment (24 for the first treatment, and the 12 periods for each of the other treatments). Thus, we have 52 independent observations per treatment in the BSs data and 84 matched-paired observations for the WSs data.[13] For BSs we use the robust rank order test Fligner and Policello (1981) for pairwise comparisons, and the Kruskal-Wallis test for comparisons of more than two categories (i.e., for $\beta_\omega = 1/2$, *4B = 2B = 0B*). For WSs data, we use Wilcoxon sign-rank test for pairwise comparisons, and the Friedman test for comparisons across all three treatments for the knife-edge case. Except for the Kruskal-Wallis and Friedman tests, we report p-values from one-sided tests since our model provides clear predictions regarding the direction of treatment effects.

For the regressions, we separately estimate on the pooled BSs data and for the WSs data, indexing

[8]Payments were converted to local currency at a rate of E\$6 = Q1 (E\$46.2 = US\$1) and participants received a show-up fee of 20Q (2.6USD).

[9]We have perfect balance regarding the order in which the within-subjects treatments were implemented across the 6 sessions.

[10]Payments were converted to local currency at a rate 40E\$ per \$1 USD, and subjects received a show-up fee \$7 USD.

[11]Although on average they acquire more information than predicted by risk-neutral model, we observe that they are prices influence their information acquisition decisions.

[12]We also estimate the probability of acquiring signals using a linear probability model. Our dependent variable is a dummy on whether an even number of balls were acquired. Our independent variables are treatment dummies, cost of acquiring information, period, period^2 and wave.

[13]Tables 3 and 4 presents the summary of results by treatment and situation, separately for each wave.

subjects by i and periods by t, the following regression model:

$$Y_{it} = \alpha_0 + \alpha_1 \beta_t^{\omega=\frac{1}{2}} 2B_i + \alpha_2 \beta_t^{\omega=\frac{1}{2}} 0B_i + \alpha_3 \beta_t^{\omega=\frac{9}{13}} 2B_i + \alpha_4 \beta_t^{\omega=\frac{9}{13}} 0B_i + \gamma Cost_{it} + X^{0}\delta + \epsilon_{it} \quad (1)$$

where our dependent variable is information acquisition, Y_{it}, either at the intensive margin (a dummy on whether any balls were acquired), or extensive margin (a discrete variable $[0, 5]$ on the number of balls acquired).[14] Our independent variables are treatment-prior dummies, and we control for the cost of acquiring information. X is a vector of controls that includes period (linear and squared), for BSs (WSs) dummy for the wave of the data (task order controls), and in some WSs specifications also individual fixed effects.

Our main hypothesis is that the offer of free information in the future increases information acquisition, $(1) Ha : \alpha_3 > \alpha_4$. However, in the rare knife-edge cases, the effect reverses as free information in the future substitutes for information acquisition, $(2) Ha : \alpha_0 < \alpha_1 < \alpha_2$

5 Results

Table 2 provides summary statistics of observed and predicted instances of "dropping out" (no information acquisition), amount of information acquired and correct guesses broken down by informational situation. Figures 1 and 2 present, for BSs and WSs, histograms of predicted and observed information acquisition by treatment and informational situation.[15] As the table and figures also illustrate, the share of observations which involve no purchase of information is lower than predicted across all informational situations. Indeed, the average number of balls purchased exceeds predictions in all situations.[16] The table as well as figures illustrates the main predictions of the model: that under asymmetric priors (i.e., $\beta_\omega = 9/13$) promising future information should induce greater information acquisition. This is especially noteworthy at the extensive margin through the reduction in the mass at zero instances in the figures, and in the table, the increase in the share of no information acquisition. However, the opposite is true in the rare knife edge case ($\beta_\omega = 1/2$): promises of future information induce lower information acquisition.

(1.1) Ha (extensive margin): For given (non-symmetric) priors ($\beta_\omega = 9/13$), subjects are more likely to acquire information when free information is promised in the future (2B, $2 : 0$) than in when no future information is offered (0B, $3 : 1$).

Results: We find that the rate of information acquisition when future information is promised is .535 (.408) for BSs (WSs), compared to .345 (0.313) when no future information is promised. Using non-parametric tests we reject the null hypothesis that the probability of acquiring information in $2B \quad 0B$ using either the BSs ($U = 2.702$, $p = 0.003$) or the WSs data $2B \quad 0B$ ($z = 2.475$, $p = 0.006$).

Our reduced form estimates confirm the results. The top panel of figure 3 plots the coefficients from a linear probability model. Appendix Table 5 presents the full results from the regressions exploring the extensive margin.[17] We reject the null hypothesis using either a linear probability model (BSs: $p = 0.004$, WSs: $p = 0.015$) or a random effects logit model (BSs: $p = 0.005$, WSs: $p = 0.009$). Using the linear

[14] We also use equation (1) to examine predictive accuracy of guesses as dependent variable.

[15] Appendix Figures 5 and 6 presents CDF's of predicted and observed information acquisition decisions for BSs and WSs.

[16] Appendix table 3 contains both predicted and observed summary statistics by treatment, conditional on the realized draws observed by subjects when they made their information acquisition decisions.

[17] For sake of brevity, in this section we report the p values for a (one sided) test of the null that $\alpha_3 > \alpha_4$, as described in equation (1).

probability model, we find that the probability of acquiring information increases by 8   17 percentage points (depending on whether we use BSs or WSs) when future information is promised.

(1.2) Ha (intensive margin): For given (non-symmetric) priors ($\beta_\omega = 9/13$), when free information is promised in the future (2B, $2:0$) subjects acquire more information compared to when no future information is promised (0B, $3:1$).

Results: With BSs (WSs) data, we find that subjects acquire on average 0.315 (0.125) more balls when future information is promised. That is an increase of 37% (16%) compared to when no future information is promised. Using non parametric tests, we reject the null hypothesis that the amount of information acquired $2B \quad 0B$ with the BSs data ($U = \quad 1.680$, $p = 0.046$). For the WSs data the result is only marginally significant ($z = 1.542$, $p = 0.062$).

Reduced form results are only marginally significant. Using the random effects Poisson model (BSs: $p = 0.065$, WSs: $p = 0.051$) or a random effects Tobit model (BSs: $p = 0.058$, WSs: $p = 0.001$).[18] Top panel of figure 4 presents coefficient plots for the random effects Poisson model. As the figure illustrates, although our results are only marginally significant, they have the predicted sign and are consistent across the sub-samples.

A further confirmation of the model is that a simple change in the parameters that modify the updated prior to a knife-edge case ($\beta_\omega = 1/2$) reverses the effects of promising future information.

(2.1) Ha (extensive margin): For symmetric knife-edge priors ($\beta_\omega = 1/2$), subjects are less likely to acquire information in 4B or 2B ($1:1$) than in 0B ($2:2$).

Results: We reject the null hypothesis that the probability of acquiring any information is equal across treatments ($\chi^2 = 13.862$, $p < 0.001$ for BSs; $Friedman = 195.2$, $p < 0.0001$; for WSs). For pairwise comparisons, we reject that $4B \quad 0B$ (BSs data: $U = \quad 3.813$, $p < 0.0001$; WSs data: $z = \quad 5.605$, $p < 0.0001$) , in $2B \quad 0B$ (BSs data: $U = \quad 1.680$, $p = 0.046$; WSs data: $z = \quad 3.953$, $p < 0.0001$) $4B \quad 2B$ (BSs data: $U = \quad 2.2253$, $p = 0.013$; WSs data: $z = \quad 3.341$, $p < 0.001$). Our reduced form models (reported in Table 5) and illustrated in the bottom panel of figure 3 also allow us to test these hypotheses and overall we find strong support for them.[19]

(2.2) Ha (intensive margin): For symmetric knife-edge priors ($\beta_\omega = 1/2$), subjects acquire less information in 4B or 2B ($1:1$) than in 0B ($2:2$).

Results: In the knife-edge cases for BSs we marginally reject ($\chi^2 = 4.752$, $p = 0.093$) and for WSs we strongly reject ($Friedman = 194.96$, $p < 0.0001$) the null hypothesis of equal information acquisition across treatments. For pairwise comparisons that information acquisition is lower when information is promised in the future, we reject using any test/data for the most extreme case ($4Bvs.0B$): $4B \quad 0B$ (BSs data: $U = \quad 2.260$, $p = 0.0119$; WSs data: $\quad 2.870$, $p = 0.002$). For our other hypotheses, although results have the predicted sign in all cases, they are only statistically significant using the (more powerful) within-subjects data: $2B \quad 0B$ (BSs data: $U = \quad 1.194$, $p = 0.116$; WSs data: $z = \quad 0.898$, $p = 0.019$); for $4B \quad 2B$, results (BSs data: $U = \quad 0.909$, $p = 0.182$; WSs data: $z = \quad 2.431$, $p < 0.007$). Our reduced form results can be seen in the bottom panel of figure 4. Using the reduced form results we see again that we consistently reject the null for the most extreme hypotheses $4B \quad 0B$ using random effects Poisson model

[18] Appendix Table 6 presents the full results for the regression estimates.

[19] The only exception is $2B \quad 0B$, where the differences are not statistically significant with the pooled BSs data, although they are significant at the $p < 0.001$ with the WSs data.

and a random effects Tobit model. We only reject the other hypotheses ($2B$ $0B$ and $4B$ $2B$) using the within-subject data.[20]

In addition to the main hypotheses, we find support for additional propositions of secondary importance. As predicted by theory, we observe no difference in the decision quality across treatments with assymetric priors. Despite differences in information acquisition by treatment, the proportion of time the subjects choose the correct jar color is no higher *2B* than *0B* (for NPT: $p = 0.368$ for BSs, $p = 0.6146$ for WSs; reduced form regressions: $p > 0.307$ for BSs and $p > 0.104$ for WSs). However in the rare knife-edge situations where even small amounts of information are valuable, providing free information in the future is expected to slightly increase the predictive accuracy of guesses. We find support for it when comparing the most extreme treatments: *0B* $(2 : 2) <$ *4B* (BSs: $U = 2.136$, $p = 0.016$, WSs: $z = 3.607$, $p < 0.001$).[21] This lends further support to the theory.

6 Conclusions and Further Research

We present results of a laboratory experiment on the impact a promise of future help may exert on individual information-acquisition effort. We observe that differences across treatments are more pronounced on the extensive than on the intensive margin: promise of delayed information makes the agents less likely to choose not to acquire any information at all. As predicted by the model, information acquired in this manner is, ex post, useless: the quality of the overall decision-making is unaffected.

Our ability to induce information acquisition through a promise of free information is, however, suggestive of possible information spillovers in group information acquisition environments with communication, such as juries: the possibility that we believe should be explored in future research. The effect we identify appears to be an important previously unobserved feature of costly attention environments and may therefore be used to identify rational inattention in the field.

In terms of institutional design, delay in information provision may be sufficient to discourage "informational drop out": situations in which agents choose to forgo attentional effort and make decisions based entirely on their prior beliefs (this appears to be a plausible explanation for certain typical features of the common law jury system). It may also be a useful tool in experimental design, as it could be used to avoid excessive drop-out by subjects that has been observed in some previous experimental studies (Elbittar et al., 2016). It might also be a feature of the registered reports –the peer-review of pre-results submissions– to encourage greater attention from reviewers.

References

Allcott, Hunt and Dmitry Taubinsky, "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 2015, *105* (8), 2501–2538.

Battaglini, Marco, Rebecca B. Morton, and Thomas R. Palfrey, "The swing voter's curse in the laboratory," *Review of Economic Studies*, 2010, *77* (1), 61–89.

[20] Appendix Table 6 presents the full results for the regression estimates.

[21] We also estimate the probability of guessing correctly the state using equation (1) with a linear probability model and a random-effects logit model. Our dependent variable is a dummy on whether they correctly guessed the color of the jar. Results are presented in Appendix Table 7. Again we find support, rejecting the one-sided hypothesis that *0B* (2 : 2) *4B*, with $p < 0.05$ for all specifications.

Bernheim, B. Douglas and Dmitry Taubinsky, "Behavioral Public Economics," in "Handbook of Behavioral Economics," Vol. 1, Elsevier B.V., 2018, chapter 5, pp. 381–516.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor, "Choose to lose: Health plan choices from a menu with dominated options," *The Quarterly Journal of Economics*, 2017, *132* (3), 1319–1372.

Bhattacharya, Sourav, John Duffy, and Sun Tak Kim, "Voting with endogenous information acquisition: Experimental evidence," *Games and Economic Behavior*, 2017, *102*, 316–338.

Blackwell, David, "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 1962, *33* (2), 719–726.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen, "Calorie posting in chain restaurants," *American Economic Journal: Economic Policy*, 2011, *3* (1), 91–128.

Caplin, Andrew and Daniel Martin, "A testable theory of imperfect perception," *Economic Journal*, 2015, *125* (582), 184–202.

_ **and** _ , "Framing as Information Design," 2018.

_ **and Mark Dean**, "Revealed preference, rational inattention, and costly information acquisition," *American Economic Review*, 2015, *105* (7), 2183–2203.

Chade, Hector and Edward E. Schlee, "Another look at the Radner-Stiglitz nonconcavity in the value of information," *Journal of Economic Theory*, 2002, *107* (2), 421–452.

Chaloupka, Frank J., Kenneth E. Warner, Daron Acemoğlu, Jonathan Gruber, Fritz Laux, Wendy Max, Joseph Newhouse, Thomas Schelling, and Jody Sindelar, "An evaluation of the FDA's analysis of the costs and benefits of the graphic warning label regulation," *Tobacco Control*, 2015, *24* (2), 112–119.

Dean, Mark and Nathaniel Neligh, "Experimental Tests of Rational Inattention," *Working Paper*, 2017, (June), 1–55.

Elbittar, Alexander, Andrei Gomberg, César Martinelli, and Thomas R. Palfrey, "Ignorance and bias in collective decisions," *Journal of Economic Behavior and Organization*, dec 2016, *In press.*

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental economics*, 2007, *10* (2), 171–178.

Fligner, Michael A and George E Policello, "Robust rank procedures for the Behrens-Fisher problem," *Journal of the American Statistical Association*, 1981, *76* (373), 162–168.

Frederick, Shane, "Cognitive reflection and decision making," *Journal of Economic perspectives*, 2005, *19* (4), 25–42.

Grosser, Jens and Michael Seebauer, "Standing versus Ad Hoc Committees with Costly Information Acquisition," *SSRN Electronic Journal*, 2018.

Guarnaschelli, Serena, Richard D. Mckelvey, and Thomas R. Palfrey, "An Experimental Study of Jury Decision Rules," *American Political Science Review*, 2018, *94* (2), 407–423.

Hannaford, Paula L., Valerie P. Hans, and G. Thomas Munsterman, "Permitting jury discussions during trial: Impact of the Arizona Reform.," *Law and Human Behavior*, 2000, *24* (3), 359–382.

Larreguy, Horacio, John Marshall, and James M Snyder, "Publicizing malfeasance: When the local media structure facilitates electoral accountability in Mexico," *The Economic Journal*, 2020.

León, Gianmarco, "Turnout, political preferences and information: Experimental evidence from Peru," *Journal of Development Economics*, 2017, *127* (December 2016), 56–71.

Radner, Roy and Joseph Stiglitz, "A Nonconcavity in the Value of Information," in M. Boyer and R.E. Kihlstrom, eds., *Bayesian models in economic theory*, Elsevier, 1984, chapter 3, pp. 33–52.

7 Tables

Table 1: Marginal value of information acquisition

Updated prior	Situation	Ball 1	Ball 2	Ball 3	Ball 4	Ball 5
$\beta_\omega = 1/2$	B_4	3.46	0	2.76	0	2.32
	B_2 (1 : 1)	4.8	0	3.46	0	2.76
	B_0 (2 : 2)	10	0	4.8	0	3.46
$\beta_\omega = 9/13$	B_2 (2 : 0)	2.22	0	2.13	0	1.92
	B_0 (3 : 1)	0	0	2.22	0	2.13
$\beta_\omega = 81/97$	B_0 (4 : 0)	0	0	0	0	0.43

Notes: Marginal value of purchasing a ball, in percentages of the prize value at the moment of information acquisition.

Table 2: Summary statistics by situation

Updated Prior	Situation	Obs.	Share No Info		Avg. Balls Purchased		Avg. Correct	
			Predicted	Observed	*Predicted*	Observed	*Predicted*	Observed
Between-subjects data								
$\beta_\omega = 1/2$	B_4	1,248	*0.749*	0.453	*0.675*	1.345	*0.706*	0.684
	B_2 (1 : 1)	581	*0.651*	0.339	*0.921*	1.551	*0.666*	0.645
	B_0 (2 : 2)	423	*0.319*	0.253	*1.324*	1.730	*0.599*	0.617
$\beta_\omega = 9/13$	B_2 (2 : 0)	667	*0.849*	0.465	*0.436*	1.168	*0.753*	0.729
	B_0 (3 : 1)	614	*0.957*	0.655	*0.212*	0.853	*0.740*	0.712
$\beta_\omega = 81/97$	B_0 (4 : 0)	211	*0.972*	0.787	*0.142*	0.559	*0.905*	0.877
Within-subjects data								
$\beta_\omega = 1/2$	B_4	1,008	*0.736*	0.546	*0.641*	1.14	*0.683*	0.676
	B_2 (1 : 1)	479	*0.647*	0.438	*0.766*	1.25	*0.612*	0.620
	B_0 (2 : 2)	361	*0.307*	0.343	*1.30*	1.50	*0.589*	0.554
$\beta_\omega = 9/13$	B_2 (2 : 0)	529	*0.853*	0.588	*0.393*	0.907	*0.704*	0.709
	B_0 (3 : 1)	490	*0.951*	0.694	*0.245*	0.782	*0.702*	0.682
$\beta_\omega = 81/97$	B_0 (4 : 0)	157	*0.987*	0.828	*0.064*	0.433	*0.847*	0.815

Notes: Summary statistics of predicted and observed actions for between- and within-subjects data, according to the updated priors generated by the informational case (situation). "Obs." denotes the number of observations collected in each informational case. "Share No Info" is the rlative number of instances where no information was acquired. "Avg. Balls Purchased" denotes the unconditional number of balls purchased. "Avg. Correct" is the share of instances of correct state of the world predictions (color of jar guesses).

8 Figures

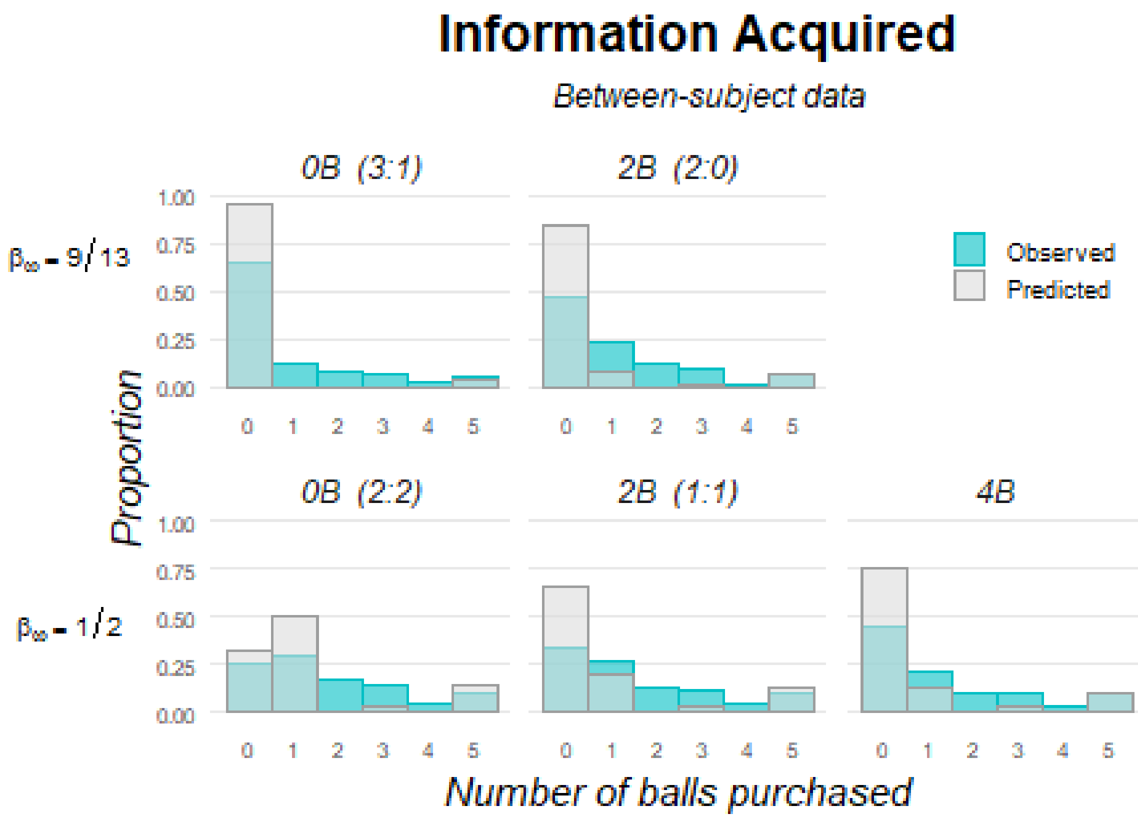


Figure 1: Histogram of predicted and observed balls purchased by treatment and informational case for between-subjects (BSs) data. Top panel for $\beta_\omega = 9/13$; bottom panel for the rare knife-edge case of indifference ($\beta_\omega = 1/2$).

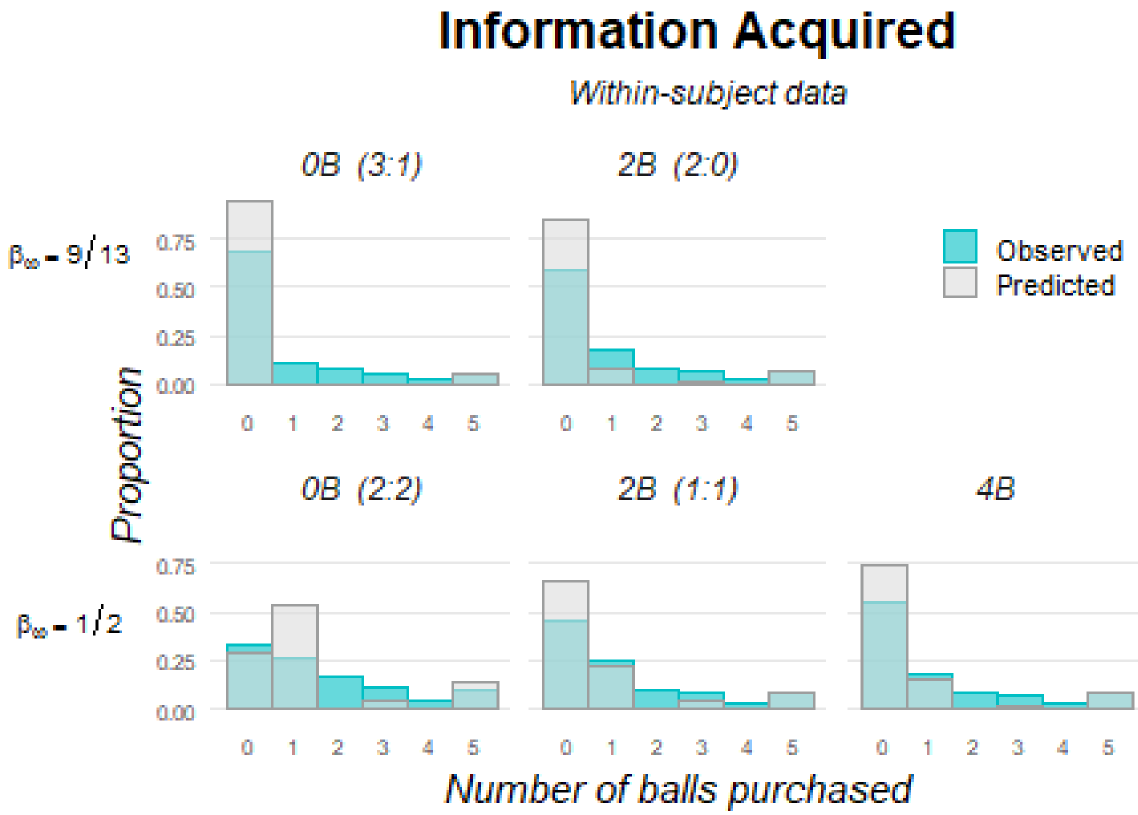


Figure 2: Histogram of predicted and observed balls purchased by treatment and informational case for within-subjects (WSs) data. Top panel for $\beta_\omega = 9/13$; bottom panel for the rare knife-edge case of indifference ($\beta_\omega = 1/2$).

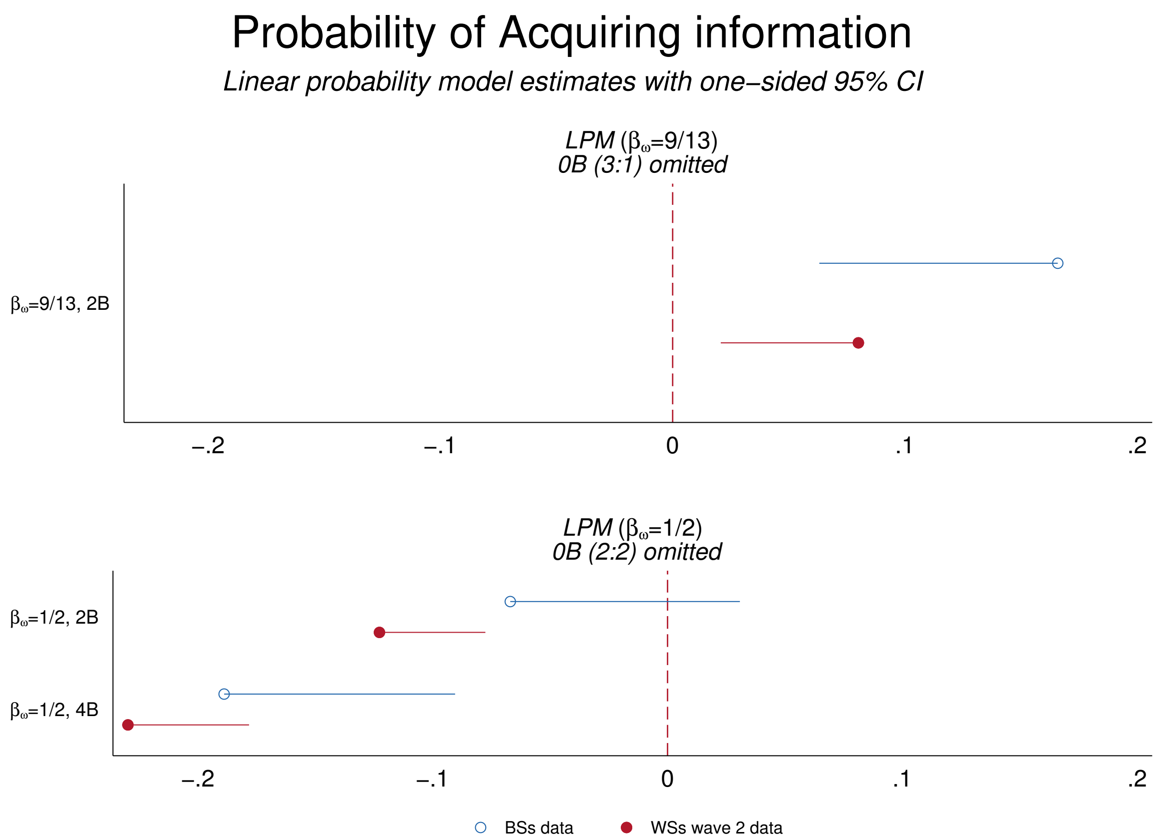


Figure 3: Coefficients plots for linear probability estimates of the effects of promising future information on the decision to acquire information (extensive margin). Results from joint estimates using all informational cases. Top panel ($\beta_\omega = 9/13$) presents estimates of promising two balls ($2B$) after information acquisition relative to the omitted category of no information ($0B$) in the future. Bottom panel presents estimates of promising two ($2B$) and four ($4B$) balls after information acquisition relative to the omitted category of no information ($0B$) in the future for the rare knife-edge case of indifference ($\beta_\omega = 1/2$).

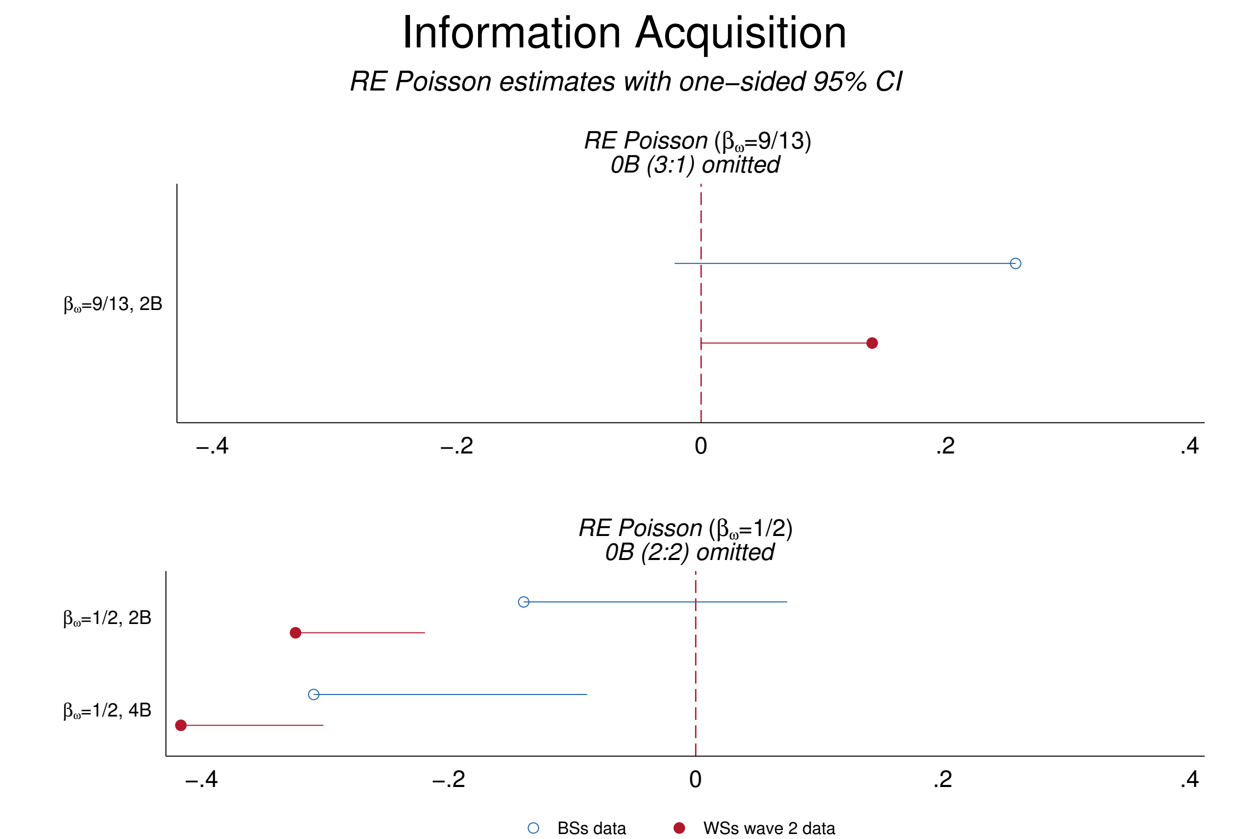


Figure 4: Coefficients plots for random effects Poisson model estimates of the effects of promising future information on information acquisition decisions (intensive margin). Results from joint estimates using all informational cases. Top panel ($\beta_\omega = 9/13$) presents estimates of promising two balls ($2B$) after information acquisition relative to the omitted category of no information ($0B$) in the future. Bottom panel presents estimates of promising two ($2B$) and four ($4B$) balls after information acquisition relative to the omitted category of no information ($0B$) in the future for the rare knife-edge case of indifference ($\beta_\omega = 1/2$).

A Additional Tables

Table 3: Summary statistics by treatment

Treatment	Subjects	Share No Info		Avg. Balls Purchased		Avg. Correct	
		Predicted	Observed	*Predicted*	Observed	*Predicted*	Observed
Between-subjects data							
B_4	52	*0.749*	0.453	*0.675*	1.35	*0.706*	0.684
B_2	52	*0.756*	0.406	*0.662*	1.35	*0.712*	0.690
B_0	52	*0.744*	0.541	*0.577*	1.101	*0.720*	0.708
Within-subjects data							
B_4	28	*0.736*	.546	*0.641*	1.14	*0.683*	.676
B_2	28	*0.755*	.517	*0.570*	1.07	*0.660*	.667
B_0	28	*0.726*	.589	*0.595*	.984	*0.684*	.657

Table 4: Summary statistics for main task by wave

Situation	No. Obs.	share no info	avg. balls	avg. correct
UFM				
B_4	576	0.340	1.679	0.724
B_2 and $1:1$	249	0.213	1.928	0.711
B_2 and $2:0$	327	0.339	1.327	0.734
B_0 and $2:2$	200	0.200	1.935	0.660
B_0 and $3:1$	279	0.577	0.993	0.767
B_0 and $4:0$	97	0.742	0.588	0.866
ESI				
B_4	672	0.549	1.058	0.650
B_2 and $1:1$	332	0.434	1.268	0.596
B_2 and $2:0$	340	0.585	1.015	0.724
B_0 and $2:2$	223	0.300	1.547	0.578
B_0 and $3:1$	335	0.719	0.737	0.666
B_0 and $4:0$	114	0.825	0.535	0.886

Notes: Summary statistics by treatment and information scenario for the main task (24 periods) of each wave.

Table 5: Information acquisition: extensive margin

	Dependent variable: Purchased any balls?			
	(1)	(2)	(3)	(4)
$\beta_\omega = \frac{1}{2}$, *2B*	0.122**	1.325**	0.107***	0.970***
	[0.054]	[0.542]	[0.025]	[0.209]
$\beta_\omega = \frac{1}{2}$, *0B*	0.189***	1.675***	0.230***	1.932**
	[0.060]	[0.577]	[0.031]	[0.277]
$\beta_\omega = \frac{9}{13}$, *2B*	-0.021	-0.169	-0.044	-0.415
	[0.061]	[0.551]	[0.032]	[0.277]
$\beta_\omega = \frac{9}{13}$, *0B*	-0.187***	-1.535***	-0.124***	-1.179***
	[0.055]	[0.466]	[0.035]	[0.327]
$\beta_\omega = \frac{81}{97}$, *0B*	-0.371***	-2.977***	-0.266***	-2.470***
	[0.059]	[0.577]	[0.046]	[0.522]
Cost	-0.006***	-0.054***	-0.007***	-0.059***
	[0.000]	[0.004]	[0.000]	[0.005]
Period	0.010**	0.080**	0.001	0.007
	[0.004]	[0.035]	[0.002]	[0.015]
Period^2	-0.000**	-0.002*	-0.000	-0.002
	[0.000]	[0.001]	[0.000]	[0.001]
Constant	0.892***	3.449***	0.726***	2.677***
	[0.050]	[0.488]	[0.029]	[0.404]
σ_u^2		1.741***		1.767***
		[0.187]		[0.221]
Individual fixed effects?	No	No	Yes	No
Wave / Task Order?	Yes	Yes	Yes	Yes
Mean of dependent variable	0.533	0.533	0.446	0.446
P-values for one-tailed Wald tests				
$\beta_\omega = \frac{9}{13}$: *2B* (2:0) = *0B* (3:1)	0.004	0.005	0.015	0.009
$\beta_\omega = \frac{1}{2}$: *2B* (1:1) = *0B* (2:2)	0.131	0.283	0.000	0.000
0B: 2:2 = 3:1	0.000	0.000	0.000	0.000
2B: 1:1 = 2:0	0.000	0.000	0.000	0.000
Number of Observations	3744	3744	4032	4032
Number of clusters	156	156	84	84
Log Likelihood	-2198.3	-1490.1	-1459.3	-1513.1
BIC	4478.8	3070.6	3001.5	3125.9
AIC	4416.6	3002.1	2938.5	3050.2

Linear probability model (1 and 3) and random effects logit model (2 and 4) estimates of the probability of information acquisition. Estimates using between- (within-) subjects data reported in columns 1 and 2 (3 and 4).Robust standard errors clustered at the individual level in brackets. Omitted treatment variable is baseline treatment *4B* ($\beta_\omega = \frac{1}{2}$): no free information before decision to acquire information.
$p < 0.10$, $p < 0.05$, $p < 0.01$

Table 6: Information acquisition: intensive margin

	Dependent variable: Number of balls purchased			
	(1)	(2)	(3)	(4)
$\beta_\omega = \frac{1}{2}$, *2B*	0.170	0.551	0.093*	0.469***
	[0.123]	[0.361]	[0.054]	[0.112]
$\beta_\omega = \frac{1}{2}$, *0B*	0.309**	1.065***	0.417***	1.229**
	[0.135]	[0.363]	[0.070]	[0.120]
$\beta_\omega = \frac{9}{13}$, *2B*	-0.112	-0.371	-0.166**	-0.446***
	[0.141]	[0.361]	[0.068]	[0.113]
$\beta_\omega = \frac{9}{13}$, *0B*	-0.370**	-0.950***	-0.306***	-0.876***
	[0.157]	[0.363]	[0.086]	[0.120]
$\beta_\omega = \frac{81}{97}$, *0B*	-0.889***	-2.096***	-0.892***	-1.953***
	[0.226]	[0.392]	[0.218]	[0.215]
Cost	-0.024***	-0.055***	-0.030***	-0.066***
	[0.001]	[0.001]	[0.002]	[0.002]
Period	0.029***	0.069***	0.005	0.002
	[0.010]	[0.020]	[0.005]	[0.007]
Period^2	-0.001**	-0.002**	-0.000	-0.001
	[0.000]	[0.001]	[0.000]	[0.001]
Constant	1.210***	3.191***	0.919***	3.031***
	[0.118]	[0.318]	[0.076]	[0.247]
$\ln(\alpha)$	-0.645		-18.850***	
	[4.044]		[4.108]	
$\ln(\sigma_u)$		1.757***		2.048***
		[0.115]		[0.172]
$\ln(\sigma_e)$		1.694***		1.823***
		[0.033]		[0.037]
Individual fixed effects?	No	No	Yes	No
Wave / Task Order?	Yes	Yes	Yes	Yes
Mean of dependent variable	1.264	1.264	1.066	1.066
P-values for one-tailed Wald tests				
$\beta_\omega = \frac{9}{13}$: *2B* (2:0) = *0B* (3:1)	0.065	0.058	0.051	0.001
$\beta_\omega = \frac{1}{2}$: *2B* (1:1) = *0B* (2:2)	0.141	0.080	0.000	0.000
0B: 2:2 = 3:1	0.000	0.000	0.000	0.000
2B: 1:1 = 2:0	0.000	0.000	0.000	0.000
Number of Observations	3744	3744	4032	4032
Number of clusters	156		84	
Log Likelihood	-4469.7	-4523.1	-3977.3	-4208.1
BIC	9029.9	9145.0	8104.0	8524.2
AIC	8961.4	9070.3	7990.6	8442.2

Random effects Poisson model (1 and 3) and random effects Tobit model (2 and 4) estimates of the amount of information acquired (number of balls purchased). Estimates using between- (within-) subjects data reported in columns 1 and 2 (3 and 4). Standard errors (clustered at the individual level for specifications 1 and 3) in brackets. Omitted treatment variable is baseline treatment *4B* ($\beta_\omega = \frac{1}{2}$): no free information before decision to acquire information. $p < 0.010$, $p < 0.05$, $p < 0.01$, $p < 0.001$

Table 7: Correct choice

	Dependent variable: Choose correct color of jar			
	(1)	(2)	(3)	(4)
$\beta_\omega = \frac{1}{2}$, *2B*	-0.035	-0.157	-0.067***	-0.257***
	[0.027]	[0.126]	[0.022]	[0.096]
$\beta_\omega = \frac{1}{2}$, *0B*	-0.067**	-0.312**	-0.122***	-0.493***
	[0.028]	[0.125]	[0.033]	[0.138]
$\beta_\omega = \frac{9}{13}$, *2B*	0.041*	0.199*	0.038	0.201*
	[0.024]	[0.120]	[0.024]	[0.111]
$\beta_\omega = \frac{9}{13}$, *0B*	0.029	0.148	0.007	0.050
	[0.023]	[0.114]	[0.021]	[0.097]
$\beta_\omega = \frac{81}{97}$, *0B*	0.188***	1.184***	0.166***	0.992***
	[0.030]	[0.250]	[0.030]	[0.204]
Cost	-0.000	-0.002*	-0.001***	-0.004***
	[0.000]	[0.001]	[0.000]	[0.001]
Period	-0.009**	-0.048**	-0.001	-0.002
	[0.004]	[0.021]	[0.001]	[0.006]
Period^2	0.000***	0.002***	0.000	0.001
	[0.000]	[0.001]	[0.000]	[0.001]
Constant	0.773***	1.250***	0.712***	0.849***
	[0.031]	[0.161]	[0.018]	[0.102]
σ_u^2		-2.430***		-3.725***
		[0.378]		[0.852]
Individual fixed effects?	No	No	Yes	No
Wave/Task order?	Yes	Yes	Yes	Yes
Mean of dependent variable	0.694	0.694	0.666	0.666
P-values for one-tailed Wald tests				
$\beta_\omega = \frac{9}{13}$: *2B* (2:0) = *0B* (3:1)	0.307	0.343	0.112	0.104
$\beta_\omega = \frac{1}{2}$: *2B* (1:1) = *0B* (2:2)	0.161	0.140	0.058	0.047
0B: 2:2 = 3:1	0.000	0.000	0.000	0.000
2B: 1:1 = 2:0	0.002	0.002	0.000	0.000
Number of Observations	3744	3744	4032	4032
Number of clusters	156	156	84	84
Log Likelihood	-2367.9	-2253.7	-2595.1	-2523.0
BIC	4818.2	4597.9	5273.2	5145.6
AIC	4755.9	4529.4	5210.2	5070.0

Linear probability model (1 and 3) and random effects logit model (2 and 4) estimates of the probability of correctly predicting the state of world (guessing the color fo the jar). Estimates using between- (within-) subjects data reported in columns 1 and 2 (3 and 4).Robust standard errors clustered at the individual level in brackets. Omitted treatment variable is baseline treatment *4B* ($\beta_\omega = \frac{1}{2}$): no free information before decision to acquire information. $p < 0.10$, $p < 0.05$, $p < 0.01$

B Additional Figures

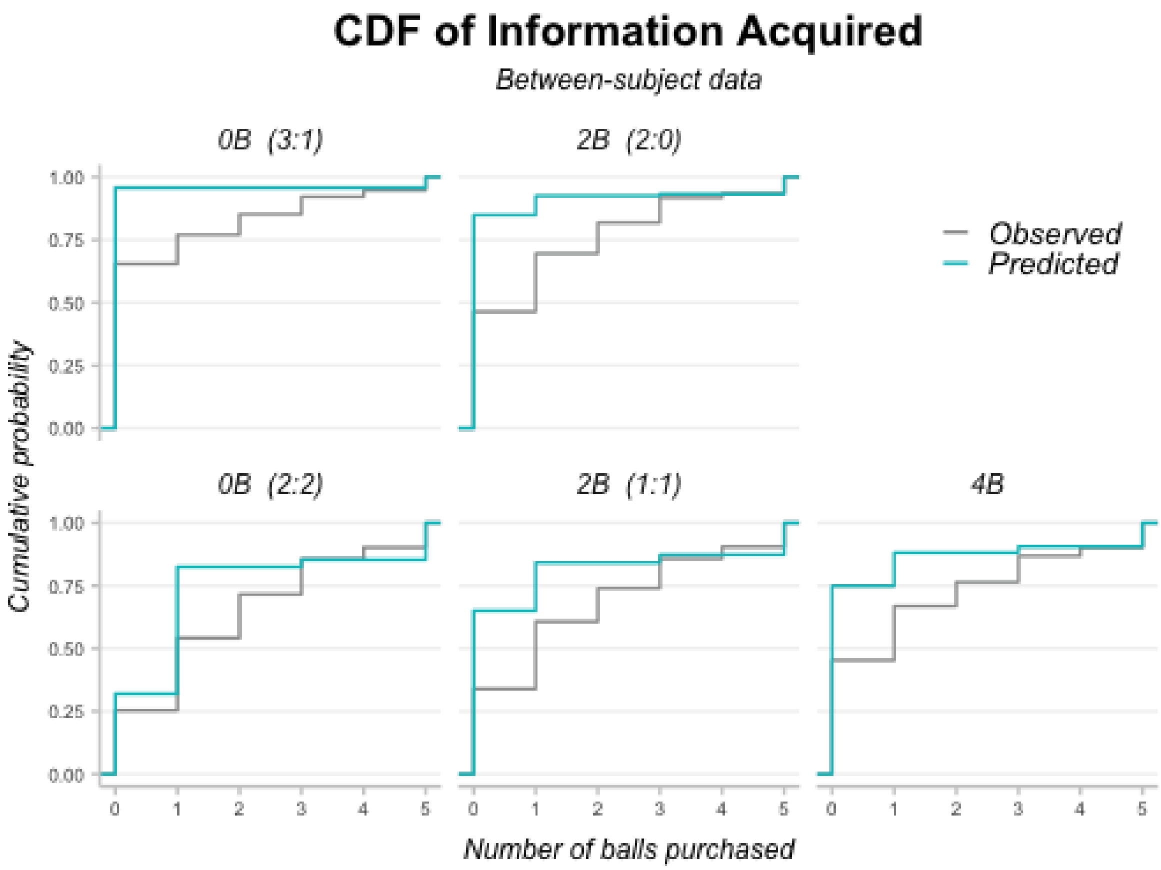


Figure 5: CDF of information purchased by individuals during last 12 periods, by treatment and information scenario for BSs

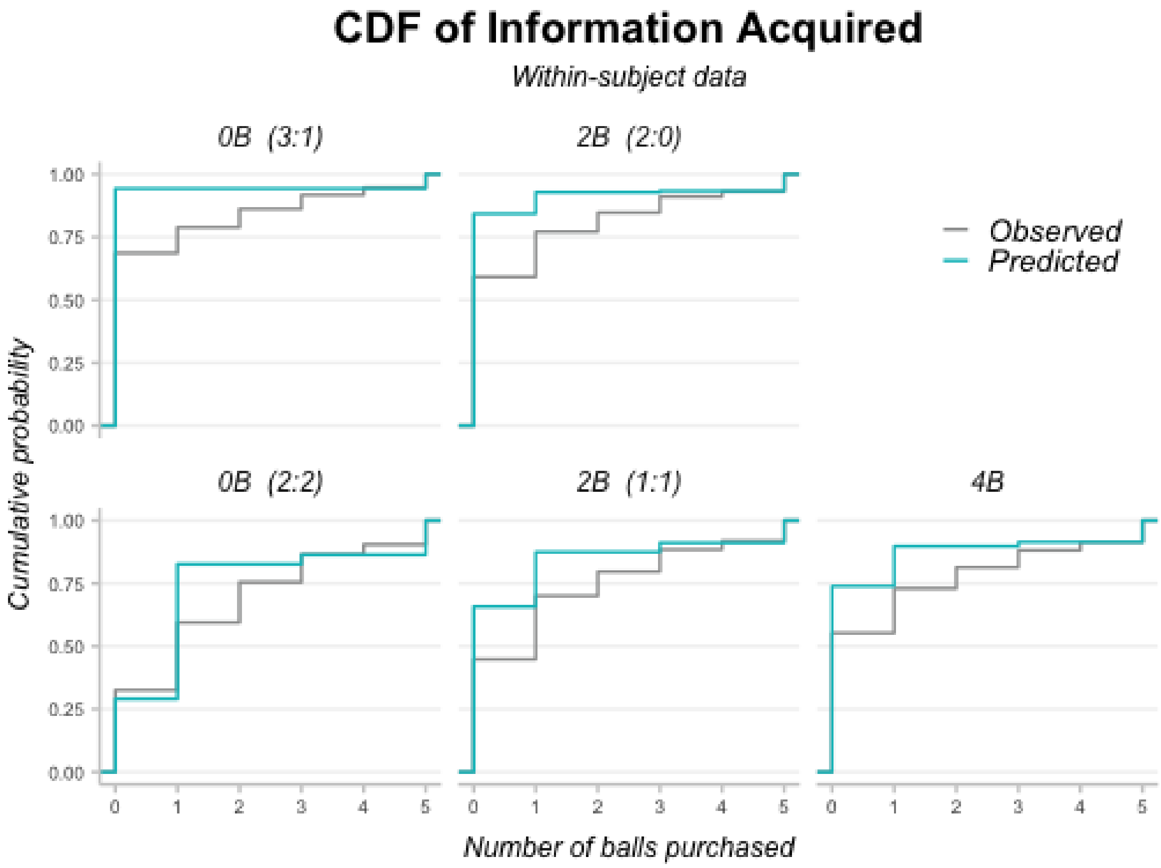


Figure 6: CDF of information purchased by individuals during last 12 periods, by treatment and information scenario for WSs

C Experimental Protocols

Upon arrival, subjects checked-in and assigned to a computer through which they interacted. Paper copies of the instructions were distributed and video of instructions with audio were played at the beginning of each experimental task. After instructions for each part and before subjects made decisions, they had to pass a test to control for comprehension of instructions. Subjects had to correctly answer all multiple-choice questions before moving on to the main task.[22] They were allowed one incorrect attempt per question before the screen locked and an experimenter was prompted. After each attempt (correct or incorrect), feedback was given to subjects to reinforce learning. For our main task, the test questions controlled understanding for: baseline probability of each state of the color of the jar (.5), conditional probability of balls being different from the true color of the jar (.4), probability of independent draws from the same color (.6), estimating earnings for guessing correctly after purchasing information, and independence of draws for the color of jars.

Across both waves, the first task consisted of 24 periods of our maininformation acquisition experimental task (for BSs data). The second task was a risk elicitation task using a multiple price list framed in the context of jars as state of the world. The third and fourth tasks were different across waves 1 and 2. In wave 1 our third (fourth) task was 12 periods of a task intended to capture independence neglect (base-rate neglect). In wave 2 sessions, we dropped the exploratory independence and base-rate neglect tasks. Instead, tasks 3 and 4 were each 12 periods of our main information acquisition task, changing treatments. In sessions 3 and 4 participants received different amounts of information after their information acquisition decision with respect to previous task(s). These tasks give rise to our within-subjects data.

Participants knew that one of the four tasks was to be randomly selected at the end, and their choice in a single (randomly selected) period for the task was to determine their earnings from the experiment. The experimental earnings was the sum of the randomly selected choice plus the show up fee and a starting balance of E$120. This was determined upon completion of the four tasks.

After completing the four tasks, participants learned their payments and took part in a post experimental survey. In the post-experimental survey we collected general demographic data (gender, age, number of siblings), unincentivized cognitive reflection test (Frederick, 2005), major and school, self reported GPA, familiarity with Bayes Theorem, number of math and stat courses taken and previous participation in research experiments.

Each session lasted about 90 min, including survey and private payment. The experiment interface was programmed using zTree (Fischbacher, 2007). Software and supporting materials are available here.

Wave 1 had between subject treatment variation, within the same session. That is, subjects within each session were randomly assigned to one of the three treatments and received video (with audio delivered via headsets) for the corresponding treatment. After pilot sessions, we discarded data from 1 session from wave 1 due to problems with the display of instruction videos in several computers that was not mentioned/revealed until the end of the session. We report data from six 12-subject sessions in wave 1.

Wave 2 had between subject treatment variation (assigned at the session level). In addition, we have within-subject data by comparing behavior of an individual in tasks 1, 3 and four. We conducted six 14-subject sessions perfectly balanced within subject treatment order.

[22] Number of questions for each task were between 3-5. Appendix D.2 shows the questions for the main task.

D Instructions

This section presents the script used to run the wave 2 sessions and the slides to present instructions.[23] Appendix section D.1 presents the instructions slides for the main task *4B* *2B* *0B* treatment. Text in black is common across all treatments. Instructions were presented through a video with audio that narrated the slides. Full video of instructions available here.

Appendix sub-section D.2 presents the control questions that followed the instructions. Appendix sub-section D.3 presents the script used to conduct wave 2 sessions.

D.1 Instruction Slides

Overview

This is an experiment about economic decision making. Various agencies have provided funds for this research. If you understand the instructions (and depending on your decisions), you can earn a considerable sum of money. At the end of today's session you will receive your earnings in cash, in private. In this experiment, the sums of money are expressed in Experimental Dollars ($E\$$). At the end of today's session, we will convert your earnings into US Dollars, at an exchange rate $E\$40 = \$1$. For today's session, you will receive an initial payment of $E\$120$.

It is important that you remain silent and not look at other people's work. If you have any questions, or need help of any kind, please raise your hand and an experimenter will come to you. If you speak aloud, you will be asked to leave the experiment. We expect and appreciate your cooperation.

Now we will describe the session in more detail.

Overview

Today's experiment consists of four parts. In each part you will make some decisions. The other participants will face similar decision-making tasks. However, their decisions will not affect your earnings and you decisions will not affect their earnings.

When all four parts have been completed, one of those parts will be randomly chosen. In the selected part one of your decisions will be chosen randomly and the result of said decision will determine your earnings.

Your final earnings for today's session will be the sum of your initial payment of $E\$120$ and your earnings in the decision which is randomly selected.

Introduction: Part 1

In the first part of the experiment there are twenty-four periods. In each period there are two jars, one red and one blue. Each jar contains 60 balls which are the same color as the jar, and 40 which are the color of the other jar.

The computer will randomly select one of the two jars. You must predict the color of the jar selected for that period. If you predict the color of the jar, you can earn up to E $\$1,000$.

To help you predict, you can observe the color of some balls from the jar. The computer will show you 4 balls for free and you will have the option to buy up to 5 additional balls (at a COST).

[23] We choose instructions for wave 2 since they are in English. Instructions for wave 1 are available from the authors upon request.

Next, we will explain this part of the experiment in m ore detail.

Jars and balls

At the beginning of each period the computer randomly selects one of two virtual jars: red or blue. The probability that the color of the jar is red is 50% and the probability that it is blue is 50%.

The color of the jar in one period does not affect the color of the jar in another period or of another participant. That is, the true color of jar is determined independently of the color of jar in another period and for other participants.

Each jar is filled with 100 virtual balls; 60 correspond to the true color of the jar and 40 are the color of the other jar. That is, the red jar contains 60 red balls and 40 blue balls. The blue jar contains 60 blue balls and 40 red balls.

In this way, the color of a ball corresponds to the true color of jar with a probability of 60%. That is, if the true color of jar is red, when drawing a ball, it will be red with 60% probability and will be blue with 40% probability. If the true color of the jar is blue, when drawing a ball, it will be blue with 60% probability and will be red with 40% probability.

Obtaining Information

Before you make your prediction about the color of the jar, you will get information by observing the colors of several balls drawn from the jar which has been selected for that period. This will work in the following way:

The computer draws a ball (chosen at random) from the jar for that period and records its color. It then deposits the ball back into the jar and all the balls are mixed. Then, another ball is randomly drawn and the color is recorded. The ball is again deposited in the jar and again the balls are mixed. This process continues until all the balls shown in the period have been drawn and recorded.

That is, each time the computer draws a ball from the jar, 60% of balls in the jar are the same color as the jar, and 40% of the balls are the other color.

Obtaining Information

Each period the computer will show you the colors of 4 balls taken from the jar at no cost. In addition, you can buy additional balls.

[You will be able to see the colors of the 4 balls that the computer will show you at no cost AFTER you decide how many balls to buy.]

[Out of the 4 balls that the computer will show you at no cost, you will be able to see the colors of 2 of these balls BEFORE you decide how many balls to buy. The colors of the other 2 balls will be revealed AFTER you decide how many you want to buy.]

[You will be able to see the colors of the 4 balls that the computer will show you at no cost BEFORE you decide how many balls to buy.]

Each additional ball you buy will have a COST. In each period, the COST of the balls will be determined randomly and will be a number between 0 and 100, all being equally probable. (The COST in one period does not affect that of other periods or other participants.)

After having seen [the colors of the 4 [2] free balls and] the COST of the additional balls, you will decide how many you want to buy, if you want to buy any.

After you decide the number of balls you want to buy, you will see the colors of the free balls and of the additional balls purchased.

Earnings from the prediction

After observing the colors of all the balls, you make your prediction regarding the color of the jar for that period. If your prediction is correct, you will earn $E\$1,000$. If you do not correctly predict the color, you will earn $E\$300$. Regardless of whether you are correct or not, you will have to pay the COST of the balls you bought that period.

That is, if your prediction about the color of the jar is correct, your earnings for the period will be given by:

$E\$1,000$ COST (# of balls purchased)

If your prediction is not correct, your earnings for the period will be given by:

$E\$300$ COST (# of balls purchased)

Feedback

During the first twelve periods, you will not be able to observe the true color of the jar at the end of each period. In the second twelve periods, you will be able to observe the true color of the jar at the end of each period.

Summary

1. In each period there are two jars and the computer will select one of these at random: the RED jar with 50% probability or the BLUE jar with 50% probability.
2. You will observe for free the colors of 4 balls. In addition, you will have the option to buy 0 to 5 additional balls, at a COST (selected randomly from between 0 to 100 for each period).
3. The color of each ball corresponds to the true color of the jar with 60% probability.
4. You must predict the color of the jar selected for that period. If your prediction is correct, you earn:

   $E\$1,000$ COST (# of balls purchased)

5. If you do not correctly predict the color of the jar, then your earnings for the period are:

   $E\$300$ COST (# of balls purchased)

D.2 Control Questions

After the instructions video ended, participants had to complete an Instructions Comprehension Test which consisted of the following (multiple choice) questions:

1. What is the probability that at the beginning of the period the computer selects the color RED jar?
2. If at the beginning of the period the computer randomly selects the BLUE jar, what is the probability that when drawing a ball from the jar it is a RED ball?
3. Assume that the jar selected at the beginning of the period is RED. Assume also that the computer has already drawn the 4 balls that it will show, and they are all red. What is the probability that if you draw an additional ball, it is RED?
4. Assume you decide to purchase three balls, at a cost of E$50 each. Also, assume you correctly predict that the jar color is BLUE. If this decision is chosen at random for your payment, what would be your profit?
5. Assume that for the last period, the true color of the jar selected by the computer was BLUE. What is the probability that for the next period the computer randomly selects the BLUE jar?

Participants had to answer all questions correctly in order to proceed to the experiment. If participants selected the correct answer for a question, they received feedback and reinforced the explanation for the correct response.

If participants selected the wrong answer, they received feedback and had a chance to answer again. If they selected an incorrect answer for a second time, the screen was locked (and requested a code that the experimental monitors had). It asked them to raise their hand so that the experimental monitor could clarify any questions or misunderstandings the subject could have.

After all participants had correctly answered all questions, they proceeded with the experimental task.

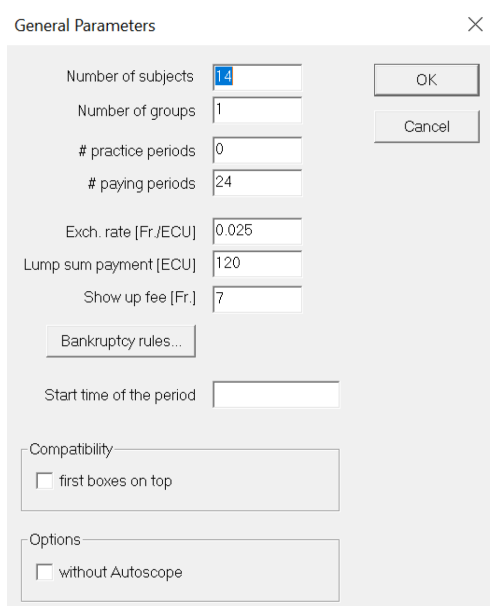
D.3 Script

Script VOI

<u>30 minutes before the session begins</u>:

1. Restart the monitor computer and the computers in the subjects' room
2. Print VOI Script, *InstrutionsComprehensionTest-Solution&Codes.pdf,* and materials for subjects:
   - Instructions VOI T#__part-1.pdf
   - Instructions _ VOI_Risk__part-2.pdf
   - RiskTask.pdf
   - VOI T# Handout part 3.pdf
   - VOI T# Handout part 4.pdf
3. Open the session # folder in the computer in the monitor room
4. Leave printed instructions <u>for part 1</u> in each subjects' computers
5. Turn on TV monitors and prepare to project the ppt "…Instructions – part 1" (Note: make sure the zTree screen is never projected; you can project from a different PC or use extend instead of duplicate mode).
   - Test to make sure volume is on at an appropriate level.
6. In the monitor computer, open zTree and open the following zTree treatment files (.ztt):
   - 1-VOI Task_english.ztt
   - 2-Risk Task_english.ztt
   - 3-VOI Task_english.ztt
   - 4-VOI Task_english.ztt
   - 5-Survey-VOI-english.ztq
7. In addition to the *client´s table*, also open the *subjects table*, *session table* and *parameters table*.
8. Change language to English in ztree treatments: *Treatment >> Language >> English* and in the zTree questionnaire (5-Survey-VOI-english.ztq) *Questionnairre >> Language >> English*

Note that *Number of subjects* need not be adjusted, unless fewer than the number of subjects recruited (14) show-up for the experiment. In that case, only the *Number of subjects* should be adjusted (and it should be adjusted in all four (4) .ztt files. No need to adjust *Number of groups.* Matching should not matter, but if need to adjust, do: *Treatment >> matching >> partner*.



Do not open the zleafs  yet in the subjects' computers.

Please make sure that these 3 JPG files are in all of the subjects' computers in *C:\Experiments\VOI*

- image_risk_options.jpg
- image_risk_JA.jpg
- image_risk_JB.jpg

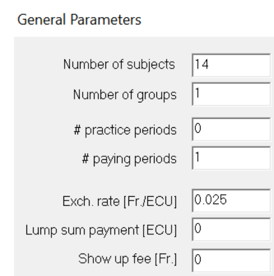
Once subjects are seated:

Read aloud to subjects: *"Welcome. Today's experiment consists of four parts. The instructions for each part will be explained through videos. Videos in the screens in front will go over the instructions. You may follow the instructions on the instruction sheets provided. At the end of the instructions, you will participate in an instructions comprehension test to ensure that everyone understands the instructions."*

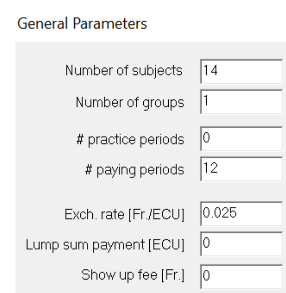
1. Start the ppt "*VOI…Instructions – part 1*"
   - When the video ends, launch the z-leafs.

2. Run ***1-VOI Task_english.ztt*** , once all subjects are connected
   - Be mindful if subjects raise their hand. If someone does raise their hand, go out (take *InstrutionsComprehensionTest-Solution&Codes.pdf*) and enter the code in their screen, make sure they understand the question they got wrong.

3. Once everyone finishes, distribute instructions for part 2 (including RiskTask.pdf handout). Start the ppt "*VOI_Risk – part 2*"
   - When the video ends, run ***2-Risk Task_english.ztt***
   - (Note: same number of subjects as in previous task, 1 group, 0 practice periods, 1 paying period, Exchange rate 0.025, lump sum payment 0, show up fee 0)
   - Be mindful if subjects raise their hand. If someone does raise their hand, go out (take *InstrutionsComprehensionTest-Solution&Codes.pdf*) and enter the code in their screen, make sure they understand the question they got wrong.



4. Distribute instructions handout for part 3. Start the ppt "*VOI…Handout – part 3*"
   - When the video ends, run ***3-VOI Task_english.ztt***
   - (Note: same number of subjects as in previous task, 1 group, 0 practice periods, 12 paying period, Exchange rate 0.025, lump sum payment 0, show up fee 0)
   - [Note that there is no instructions comprehension test for this part]



5. Distribute instructions handout for part 4. Start the ppt "*VOI…Handout – part 4*"
   - When the video ends, run ***4-VOI Task_english.ztt***
   - (Note: same number of subjects as in previous task, 1 group, 0 practice periods, 12 paying period, Exchange rate 0.025, lump sum payment 0, show up fee 0)
   - [Note that there is no instructions comprehension test for this part]

   Read aloud to subjects: *"The experiment is over. Thank you for your participation. While we prepare your payments, we ask that you complete a short questionnaire."*

6. Run ***5-Survey-VOI-english.ztq*** (while subjects complete the questionnaire, prepare to pay subjects.)
   - The amount to pay each subject is in the $MoneyToPay$ variable, in the *session table*.
   - The zTree generated pay file ($*.pay$) will have the corresponding payment to each subject, with the subject name.

7. Pay subjects

Once subjects have left:

1. Close zleafs. Close (and save) zTree .ztt files.
    - Put all data and zTree files in the *data* folder in the corresponding session.

2. Record log file for the session.