

# Does Free Information Provision Crowd Out Costly Information Acquisition? It's a Matter of Timing

## Authors:

Diego Aycinena<sup>a</sup>

Alexander Elbittar<sup>b</sup>

Andrei Gomberg<sup>c</sup>

Lucas Rentschler<sup>d</sup>

January 2023

Working Paper

The Center for Growth and Opportunity at Utah State University is a university-based academic research center that explores the scientific foundations of the interactions between individuals, business, and government.

This working paper represents scientific research that is intended for submission to an academic journal. The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Center for Growth and Opportunity at Utah State University or the views of Utah State University.

<sup>a</sup> Department of Economics at Universidad del Rosario and Economic Science Institute Chapman University, [diego.aycinena@urosario.edu.co](mailto:diego.aycinena@urosario.edu.co)

<sup>b</sup> Department of Economics at Centro de Investigación y Docencia Económicas (CIDE), [alexander.elbittar@cide.edu](mailto:alexander.elbittar@cide.edu)

<sup>c</sup> Department of Economics at Instituto Tecnológico Autónomo de México (ITAM), [gomberg@itam.mx](mailto:gomberg@itam.mx)

<sup>d</sup> Department of Economics and Finance at Utah State University and the Center for Growth and Opportunity, [lucas.rentschler@usu.edu](mailto:lucas.rentschler@usu.edu)



The Center for  
**Growth and Opportunity**  
at Utah State University

# Abstract

Conventional wisdom suggests that promising free information to an agent would crowd out costly information acquisition. We theoretically demonstrate that this intuition only holds as a knife-edge case in which priors are symmetric. Indeed, when priors are asymmetric, a promise of free information in the future induces agents to increase information acquisition. In the lab, we test whether such crowding out occurs for both symmetric and asymmetric priors. Our results are qualitatively in line with the predictions: when priors are asymmetric, the promise of future free information induces subjects to acquire more costly information.

**Keywords:** information acquisition, rational ignorance, experiments

**JEL Classification Numbers:** C44, C91, D80

\*For helpful comments, we gratefully acknowledge Sourav Bhattacharya, Eduardo Ferraz, César Mantilla, and Santiago Sautua as well as participants at the 2018 North American Economics Science Association Conference in Antigua, the 2018 Meetings of the Society for Social Choice and Welfare, the 2018 Utah Experimental Economics Conference, the 2019 Bogotá Experimental Economics Conference, and 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 generous financial support from the Economic Science Institute at Chapman University. Elbittar would like to acknowledge generous financial support from CIDE. For the experimental sessions conducted in Guatemala (wave 1), *Acuerdo Ministerial SP-M-466-2007* (regulating human clinical trials in Guatemala) did not apply to our study, and at the time, no IRB existed at any of the authors' institutions. Nevertheless, we obtained informed written consent from all participants and adhered to standard experimental protocols (no deception was used, and we ensured privacy and security of the data). The experimental sessions conducted at the Economic Science Institute (wave 2) were covered under IRB approval #0910H017. An early working paper version circulated under the title of "Rational Inattention and Timing of Information Provision."

# 1 Introduction

Providing easily digestible information to “nudge” individual behavior has recently become a major theme in the literature on consumer choice.<sup>1</sup> It has also been demonstrated to play an important role in improving political choices.<sup>2</sup> An underexplored issue is the impact that such exogenous information provision may have on incentives for costly information acquisition that individuals could undertake on their own. At first glance, it may appear that the promise of free information would simply crowd out costly information acquisition. However, it turns out that for this to be true, the agent must be indifferent between actions so that an arbitrarily small amount of information would be decisive. Since this relies on initial exact indifference, it would not be something one could expect to observe often in real life.

In the absence of exact indifference, the promise of free information in the future can incentivize the agent to acquire *more* information than they would otherwise. The intuition for this result is grounded in the fact that information that does not affect one’s choice has no (instrumental) value. Unless one is ex ante indifferent, a small amount of information will do little to change posterior beliefs and will therefore not affect choices. That is, weak signals do not affect optimal behavior and are payoff irrelevant ex post, no matter their realization. This gives rise to the Radner and Stiglitz (1984) observation that the marginal value of small amounts of information is almost always zero. Thus, individual demand for information would normally exhibit discontinuities at zero since there is a minimum scale at which information should be acquired. This is the well-known Radner-Stiglitz nonconcavity in the value of information.<sup>3</sup>

A promise of delivering free additional information *after* an agent completes her costly acquisition can “smooth out” this nonconcavity. Indeed, if an agent knows that additional information will be provided, any information she acquires on her own will be valuable with positive probability so that the marginal value of a small amount of information is also positive. Hence, the promise of free information can stimulate acquiring small amounts of information, which otherwise should never be acquired.

In this paper, we provide a model of costly information acquisition in which agents are promised free information in the future, and test the predictions of this model in the lab. In our experimental design, subjects must choose between two options that are equally likely to be correct and can purchase information relevant to this decision. They also know that some relevant information will be provided for free. We vary the timing with which this free information is observed. In particular, the free information is provided before or after the information acquisition decision, or half is provided before and half after. Since free information that is observed before individual information acquisition can lead to updated priors, our design features both symmetric priors and asymmetric priors.

---

<sup>1</sup>Consider, for instance, Bollinger et al. (2011) on calories for Starbucks purchases, Chaloupka et al. (2015) on smoking warning labels, Bhargava et al. (2017) on improvements in the presentation of health insurance plans, and Allcott and Taubinsky (2015) on comparisons of light bulb technologies. For a recent review, see Bernheim and Taubinsky (2018).

<sup>2</sup>For example, León (2017) studies the impact of information on monetary penalties for abstention on turnout in Peruvian municipal elections. Larreguy et al. (2020) provide evidence regarding the impact of information about government performance on voting outcomes in Mexico.

<sup>3</sup>Chade and Schlee (2002) show that the Radner-Stiglitz nonconcavity is an extremely robust feature of costly information acquisition environments.

We conducted two waves of data collection. In the first, the timing of information acquisition was varied between subjects, while in the second, it was varied within subjects. Our experimental design across two waves allows us to explore the predictions of the model using both between- and within-subject variation. The results, both at the aggregate and the individual level, are qualitatively consistent with the predictions of the model: we find that when priors are asymmetric, subjects acquire more information when information is promised in the future. However, when priors are symmetric, the opposite is true.

The potential complementarity between free information and incentives for information acquisition has not, to the best of our knowledge, been clearly stated or tested. However, our work relates 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). Similar to these studies, we explore the consequences of agents rationally deciding how much costly information to acquire. Promising additional information in the future allows us to directly vary the set of available informational strategies. The interaction between individual information acquisition and the timing of free information provision that we explore here may be seen as a novel implication of rational inattention.<sup>4</sup> Our experimental test also relates to recent laboratory studies on costly information acquisition in collective decision-making (Bhattacharya et al., 2017; Elbittar et al., 2020; Grosser and Seebauer, 2016).<sup>5</sup>

## 2 Theory

We consider a simple model of information acquisition with two possible states of the world,  $\Omega = \{R, B\}$ , and two possible actions,  $a \in \{r, b\}$ .<sup>6</sup> A risk-neutral agent, who assigns a prior probability  $P(R) = \beta \in (0, 1)$  to  $\omega = R$ , chooses action  $a$ , and her utility,  $U$ , depends only on  $a$  and  $\omega$ .<sup>7</sup>

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

Each signal  $S_i \in \{\hat{r}, \hat{b}\}$  is binary and correlated with the state of the world so that  $P(S_i = \hat{r} | \omega = R) = P(S_i = \hat{b} | \omega = B) = p \in (\frac{1}{2}, 1)$ . Conditional on the state, signals are independent and identically distributed (i.i.d). Before choosing  $a$ , the agent decides how many signals she would like to

<sup>4</sup>A somewhat similar effect has been noted by Caplin and Martin (2018), who explore, from a rational inattention standpoint, how varying default options presented to subjects may nudge them to either acquire information or to “drop out.”

<sup>5</sup>These studies establish a framework in which potential informational spillovers from increased information acquisition generated by delayed communication between group members could be explored. It is of interest that, in the last two studies referenced, the subjects consistently acquired less information than predicted by the experimental setting. Elbittar et al. (2020) propose that this may arise from biased priors, which effectively lead to a lack of information acquisition of the sort we study in this current paper. An offer of future information, which could be interpreted as arising from the promise of jury deliberation, would be expected to help resolve this problem.

<sup>6</sup>This simplified setting matches the model tested in our experiment, but the results we present are fairly general and follow from Radner and Stiglitz (1984) and Chade and Schlee (2002).

<sup>7</sup>We assume risk neutrality in this section for simplicity. Introducing risk aversion does not qualitatively change predictions. In particular, individual willingness to pay for information is nonmonotonic in the level of risk aversion (assuming, for instance, a constant relative risk aversion (CRRA) utility function), with effects of risk aversion being exceedingly small for any plausible level.

purchase,  $v$ , at a constant per-signal cost  $c > 0$ . She also knows that a fixed number of free signals,  $n$ , will be observed after she chooses how many signals to purchase but before she chooses  $a$ . We concentrate on the interaction between the timing at which the free signals are observed by the agent and her decision to acquire additional signals.

The posterior belief of a Bayesian agent who observes  $m$  signals  $S_i = \hat{r}$  and  $k < m$  signals  $S_i = \hat{b}$  is

$$P\left(R|m \text{ signals } \hat{r} \text{ and } k < m \text{ signals } \hat{b}\right) = \frac{\beta p^{m-k}}{\beta p^{m-k} + (1-\beta)(1-p)^{m-k}}.$$

Since utility  $U$  depends only on  $a$  and  $\omega$ , information only has an instrumental value. It is only ex ante valuable if it has the potential to change an agents action,  $a$ . The agent changes an action when  $E[U(\omega, r)|\text{signals}] > E[U(\omega, b)|\text{signals}]$ .<sup>8</sup> Thus, the smallest number of signals such that observing them has ex ante positive value is

$$n^* = \min \left\{ k \in \mathbb{N} : \frac{\beta p^n}{\beta p^n + (1-\beta)(1-p)^k} > q > \frac{\beta(1-p)^n}{\beta(1-p)^n + (1-\beta)p^n} \right\}.$$

As long as  $\beta \neq q$  and  $p$  is close enough to  $\frac{1}{2}$  (i.e., information comes in small increments),  $n^* > 1$ : small amounts of information are useless. This is the manifestation of the Radner-Stiglitz nonconcavity in our setting. On the other hand, as long as the agent observes  $N \geq n^* - 1$  signals in addition to whatever information she acquires, even a single signal will always have a positive value. Thus, provided  $c$  is sufficiently small, the promise of free signals will increase the level of costly information acquisition. Since acquiring a single signal is useless without the promise of free signals, the free signals cannot displace the agent's own information acquisition effort; it can only be complementary to it.

There is, however, a special case in which any small amount of information is ex ante valuable. This happens if  $\beta = q$  so that without observing any signals, the agent is indifferent between the two actions. In this case, standard Bayesian updating implies that she would choose  $a = r$  whenever  $m \equiv \#\{i : S_i = \hat{r}\}$  is bigger than  $k \equiv \#\{i : S_i = \hat{b}\}$  and would choose  $a = b$  whenever  $k > m$  (the agent is indifferent whenever  $k = m$ ). If  $\beta = q$ , then 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 observing even a single signal would break the indifference so that the marginal value of information at zero is, in fact, positive. This leads to a reversal of the expected impact of promising free signals: there is no nonconcavity in the value of purchased signals to smooth via the promise of free signals, and the agent is predicted to “free ride” on this promise.

Except in the knife-edge case outlined above, the promise of  $n$  free signals can induce greater information acquisition. While we are primarily interested in testing this result, it is worth noting that the increased information acquisition will not improve the expected quality of the agent's decisions. To see this, suppose the  $n$  free signals,  $S_1^f, \dots, S_n^f$ , were observed before the agent decides how many signals to purchase at cost. Defining the difference between the number of free signals that indicate  $R$  and  $B$  as

<sup>8</sup>This is the same as  $-(1-P)q > -(1-Q)P$ , or  $\pi > q$  (where  $\pi$  is the posterior probability of the first equation).

$X^f = \#\{i : S_i^f = \hat{r}\} - \#\{i : S_i^f = \hat{b}\}$ , we observe that in order for the realized free signals to impact the agent’s choice of  $v$ , she would have to purchase more signals than  $X^f$ .

Indeed, suppose the number of signals she purchases is weakly less than  $|X^f|$ . Then, even if all of the signal realizations are identical, the sign of  $m - k$  is the same as the sign of  $X^f$ , implying the same choice of  $v$ . Thus, she should either buy a lot of information or none—Radner and Stiglitz (1984) in action. Of course, even a single signal would be valuable if the realizations of the free signals “tie” so that  $X^f = 0$ . 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  $P(X^f = 0) = \frac{n!}{(\frac{n}{2}!)^2} p^{\frac{n}{2}} (1 - p)^{\frac{n}{2}} > 0$ .

### 3 Experimental Design

Our main experimental task follows a standard information acquisition environment (Elbittar et al., 2020; Guarnaschelli et al., 2018; Battaglini et al., 2010). In each of the 24 periods, subjects earn more money if they correctly guess the true binary state of the world,  $\Omega = \{R, B\}$ , framed as guessing the color of a jar. Subjects know that the true color of the jar is equally likely to be red or blue. The jar contains 100 balls: 60 corresponding to the jar’s color and 40 of the other color. Each signal is a ball randomly drawn from the jar (with replacement) so that the color of a signal corresponds to the state with probability 0.6. Subjects know they will receive four free signals and can acquire up to five additional signals at a constant marginal cost  $c$ . After all signals are observed, the subject guesses the color of the jar.

In each period, the cost of purchased signals is drawn from a uniform distribution between 0 and 100. Subjects observe the realized cost before deciding how many signals to purchase. After observing the realized signals, they proceed to guess the color of the jar. Those who guess correctly earn E\$1,000 minus the cost of purchased signals; otherwise, they earn E\$300 minus the cost of purchased signals.<sup>9</sup> Note that we deliberately induce payoffs and initial priors so that in the absence of any signals, subjects are exactly indifferent between the jars. This made a subject’s guess extremely simple: a Bayesian subject would simply guess the jar whose color corresponded to the majority of the observed signals. Importantly, and as we demonstrate below, subjects overwhelmingly adopted this heuristic. This is an asset to our design as it allows us to focus on the information acquisition decision, which is our primary outcome of interest.

Our treatment variable is the timing with which the realized free signals are observed. In treatment  $0B$ , no free signals are observed after the information acquisition decision. That is, all four free signals are observed before the decision to acquire information. In treatment  $2B$ , subjects observe two of the free signals before the information acquisition decision. Thus, they are promised two signals in the future. Finally, in treatment  $4B$ , subjects know they will see all four free signals after the decision to acquire information.

A feature of our design is that, in addition to variation in the timing of information provision, we have exogenous variation in the realizations of the free signals that are observed before the information acquisition decision. Since subjects will update their priors based on the realizations of any free signals they

<sup>9</sup>In the first 12 periods, subjects do not learn the true state before proceeding to the next period. In the second 12 periods, they do. Instructions are available in appendix D.

observe before deciding how many to purchase, the variation in these realizations gives rise to different situations—updated priors  $\beta_\omega = \{1/2, 9/13, 81/97\}$  with different corresponding predictions. We will refer to a situation within a treatment by noting the number of realized signals that corresponded to a single color. For example,  $0B(4 : 0)$  refers to the situation in which a subject observes all four free signals before deciding how many to purchase and all four free signals are the same color. Due to the nonconcavity in the value of information, the main prediction of the model is that promising future information—as in  $2B$  relative to  $0B$ —should induce greater information acquisition when the updated prior is asymmetric ( $\beta_\omega = 9/13$ ). The opposite is true in the knife-edge case, where the updated prior leads to indifference ( $\beta_\omega = 1/2$ ).<sup>10</sup>

The intuition is as follows: consider situation  $0B(3 : 1)$ , where a subject observes four free signals before she decides how many to acquire at cost and is shown three signals of the same color ( $\beta_\omega = 9/13$ ). Purchasing two signals or less cannot provide evidence sufficient to change the optimal guess. Thus, she would have to purchase at least three signals at cost for this additional information to have any value. Such a subject will drop out by purchasing no signals at all unless the realized cost is extremely low. Contrast this with  $2B(2 : 0)$ , where a subject observes both signals of the same color ( $\beta_\omega = 9/13$ ). In this case, purchasing a single additional signal has a positive value since the promised information (two additional free signals) can smooth out the nonconcavity.

Table 1 presents the predicted marginal value of each additional signal across every informational situation in which a risk-neutral subject may find herself.<sup>11</sup> The different predictions across informational situations are reflected in the marginal value of additional signals.

### 3.1 Experimental Tasks and Protocols

Upon arrival, subjects checked in and were assigned to a computer. For each of the four experimental tasks, paper copies of the instructions were distributed and videos of the instructions were played.<sup>12</sup> Before subjects made decisions, they had to pass a quiz to ensure their comprehension.

In all sessions, the first task subjects faced was the 24 periods of the jar-guessing task described above. The second task was a risk elicitation task using a multiple-price list format. The two remaining tasks were additional jar-guessing tasks (of 12 periods each). Subjects knew that one of the four tasks would be randomly selected and their decisions in a single (randomly selected) period of the selected task would determine their earnings.<sup>13</sup> Total earnings were the sum of the earnings from the randomly selected choice plus the show-up fee and a starting balance of E\$120.

<sup>10</sup>An additional advantage of this set-up is the effective oversampling of the otherwise rare knife-edge case ( $\beta_\omega = 1/2$ ).

<sup>11</sup>We use the risk-neutral benchmark for two reasons: tractability and predictive power. Introducing risk aversion into the model has a nonmonotonic effect in the level of risk aversion. Further, even when considering a level of risk aversion that maximizes the change in predicted information acquisition, the effect is negligible.

<sup>12</sup>Headphones were provided when treatments were varied between subjects within a session.

<sup>13</sup>If the risk elicitation task was chosen, one of the 10 lottery choices was randomly chosen for payment.

We conducted two waves of sessions.<sup>14</sup> In the first wave, 72 undergraduates (mainly) from Universidad Francisco Marroquín (UFM) participated in the experiment.<sup>15</sup> Within each session, a third of the subjects were randomly assigned to one of the three treatments for the first task; that is, we implemented all treatments within the same session (to different participants). We used the same set of (predefined random) draws across treatments. The third (fourth) task was 12 periods of a task intended to measure independence neglect (base rate neglect). We report data from six 12-participant sessions.<sup>16</sup>

In the second wave, 84 undergraduates from Chapman University took part.<sup>17</sup> In each of these sessions, 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, and this treatment was varied between sessions. In addition, for tasks three and four, we repeated the main jar-guessing task for both of the other two treatments. We report data from six 14-subject sessions.<sup>18</sup> Thus, in these sessions we have a between-subject variation task (in the first 24-period) that is comparable with the first wave and within-subject variation from tasks one, three, and four.

After completing the four tasks, subjects took part in a postexperimental survey. We collected general demographic data (gender, age, number of siblings), major and school, self-reported GPA, familiarity with Bayes's rule, number of courses taken across different topics (math, statistics, and ethics), and previous participation in research experiments. We also conducted an unincentivized cognitive reflection test (Frederick, 2005). Each session lasted about 90 minutes, including the survey and private payment. The experiment interface was programmed using zTree (Fischbacher, 2007).

---

<sup>14</sup>The motivation for the two waves of data collection is as follows. First, the results from the first wave of data collection were so striking that we wanted to replicate our results to ensure their robustness. Second, we wanted to ensure that our results were not driven by a particular subject pool, so we ran the second wave of sessions at another laboratory (further, with one subject pool in Guatemala and the second from the United States, we can be confident that our data do not contribute to the WEIRD problem Kanazawa, 2020). Third, we varied treatments between subjects and within subjects in the second wave of data collection to measure individual-level treatment effects.

<sup>15</sup>Payments were converted to local currency at a rate of  $E\$6 = Q1$  ( $E\$46.2 = 1$  USD), and participants received a show-up fee of  $Q20$  (2.6 USD).

<sup>16</sup>We discarded data from one session from this first wave due to problems with the display of instruction videos on several computers, which was not revealed until the end of the session.

<sup>17</sup>Payments were converted to local currency at a rate  $E\$40$  per 1 USD, and subjects received a show-up fee 7 USD.

<sup>18</sup>We have perfect balance regarding the order in which the within-subject variation of treatments was implemented across the six sessions.



## 4 Data and Empirical Strategy

We exploit both the between- and within-subject variation in our data. We pool data from the 24 periods of our main task across the two waves of data collection and refer to this as the between-subject (BSs) data.<sup>19</sup> In addition, from our second wave of sessions, we compare the decisions of the 84 subjects in the first task (which corresponds to one of our three treatments) to their decisions in tasks three and four (which correspond to the other two treatments). This is the within-subject (WSs) data.<sup>20</sup>

We test our main hypotheses using both nonparametric tests and reduced-form regressions. For the nonparametric tests, we take the individual-level average within a treatment to be an independent observation. Thus, we have 52 independent observations per treatment in the BSs data and 84 matched-paired observations in the WSs data.<sup>21</sup> For the BSs data, 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  $\beta_\omega = 1/2$ ,  $4B = 2B = 0B$ ).<sup>22</sup> For the WSs data, we use the Wilcoxon sign-rank test for pairwise comparisons and the Friedman test for comparisons of more than two categories. 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.

To evaluate our main hypotheses using regressions, we separately report estimates obtained using the BSs data or the WSs data. Our model is

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

where  $Y_{it}$  is either information acquisition at the intensive margin (whether any signals were purchased) or at the extensive margin (the number of signals purchased).<sup>23</sup> Our independent variables are treatment and prior dummies, and we control for the cost of acquiring information.  $X$  is a vector of additional controls that includes period (linear and squared), a dummy for the wave of the data (task order controls) for BSs (WSs) data, and, in some WSs specifications, individual fixed effects. In the interest of brevity, we summarize regression results using coefficient plots and relegate tables of the relevant regressions to the appendix.

If the promise of free signals in the future increases information acquisition when priors are asymmetric, then  $\alpha_3 > \alpha_4$ . Further, if this same promise reduces information when priors are symmetric, then  $\alpha_0 < \alpha_1 < \alpha_2$ .

<sup>19</sup>Our results are robust to considering each sample separately. This is illustrated in figure B.1, which contains coefficient plots from regressions estimating the probability of purchasing signals, with the results estimated separately for each sample. Reported coefficients correspond to situations, broken down by prior. Figure B.2 contains similar coefficient plots where the dependent variable in question is the number of signals purchased.

<sup>20</sup>Note that there is perfect balance in the WSs data regarding the order in which the three treatments were run. While the within-subject variation present in these data yields more statistical power, this was not our motivation for varying treatments on a within-subject basis. Rather, this variation allows us to measure individual-level treatment effects.

<sup>21</sup>Tables A.1 and A.2 present the summary of results by treatment and situation, separately for each wave.

<sup>22</sup>Feltoch (2003) notes that the “robust rank-order test is a modification of the Wilcoxon–Mann–Whitney test, designed to be appropriate in more situations than Wilcoxon–Mann–Whitney.”

<sup>23</sup>We also report regressions where the dependent variable is the accuracy of guesses.

## 5 Results

Before turning to the treatment differences we are primarily interested in, we consider a number of ancillary predictions as a check on whether subjects understood the environment. First, recall that a Bayesian subject would guess the color that corresponded to the majority of observed signals (and would be indifferent in case of a tie). Decisions in our experiment are consistent with this rule 94.4 percent (96.5 percent) of the time in the BSs (WSs) data. Second, as table 1 illustrates, an idiosyncratic feature of our experimental design is that purchasing a positive and even number of signals is dominated since the last of these would never make the subject strictly prefer to change her decision. We observe that, conditional on purchasing information, subjects purchase an odd number of signals 73.2 percent of the time (77.6 percent during the second half).

Table A.2 provides summary statistics of observed and predicted instances of “dropping out” (no information acquisition), amount of information acquired, and proportion of correct guesses by treatment and informational situation. Figure 1 presents, separately for the between- and the within-subject data, histograms of predicted and observed information acquisition by treatment and informational situation.<sup>24</sup> We observe that, across the informational situations, subjects purchase information more frequently than predicted and also purchase more signals than predicted.<sup>25</sup> At the same time, the data strongly support our main theoretical predictions. Thus, under asymmetric priors (i.e.,  $\beta_\omega = 9/13$  in top row), a promise of future information induces greater information acquisition. This effect is strongest on the extensive margin: subjects drop out less and acquire at least some information more frequently. As predicted, the opposite is true in the knife-edge case ( $\beta_\omega = 1/2$ ), where a promise of future information induces lower information acquisition.

**Result 1.1 (extensive margin):** *With asymmetric priors ( $\beta_\omega = 9/13$ ), subjects are more likely to acquire a positive amount of information when free information is promised in the future ( $2B(2 : 0)$ ) than when no future information is offered ( $0B(3 : 1)$ ).*

**Support:** *The rate of information acquisition when free signals are promised after the purchasing decision is 0.535 (0.408) for the BSs (WSs) data and only 0.345 (0.313) when no free signals will be observed in the future. Using nonparametric tests, we reject the null that the probability of acquiring information in  $2B(2 : 0) \leq 0B(3 : 1)$  using either the BSs data ( $U = 2.702$ ,  $p = 0.003$ ) or the WSs data  $2B(2 : 0) \leq 0B(3 : 1)$  ( $z = 2.475$ ,  $p = 0.006$ ).*

*Our reduced-form estimates confirm the result. The top panel of figure 2 plots the relevant coefficients from a linear probability model. The full results of the regressions exploring the extensive margin are reported in table A.3.<sup>26</sup> We reject the null hypothesis using either a linear probability model (BSs:  $p = 0.004$ , WSs:  $p = 0.015$ ) or a random*

<sup>24</sup>Appendix figures B.3 and B.4 present cumulative density functions (CDFs) of predicted and observed information acquisition decisions for BSs and WSs.

<sup>25</sup>Appendix table A.1 contains both predicted and observed summary statistics by treatment (not broken down by situation), conditional on the realized draws observed by subjects when they made their information acquisition decisions.

<sup>26</sup>For sake of brevity, in this section we report the  $p$  values for a (one-sided) test of the alternative hypothesis that  $\alpha_3 > \alpha_4$ , as described in equation (1).

effects logit model (BSs:  $p = 0.005$ , WSs:  $p = 0.009$ ). Using the linear probability model, we find that the probability of acquiring information increases by 8–17 percentage points (depending on whether we use BSs or WSs data) when future information is promised.

Having established that the promise of free signals increases the probability that a subject will purchase signals when priors are asymmetric, we turn to the intensive margin. Does the promise of free signals lead to more signals being purchased when priors are asymmetric?

**Result 1.2 (intensive margin):** *When priors are asymmetric ( $\beta_\omega = 9/13$ ), the promise of free information in the future ( $2B(2 : 0)$ ) leads subjects to acquire more information at cost, relative to the situation in which no free signals are promised in the future ( $0B(3 : 1)$ ).*

**Support:** *In the BSs (WSs) data, subjects acquire 0.315 (0.125) more signals when future information is promised, on average. This is an increase of 37 percent (16 percent) compared to the situation when no future information is promised. Using nonparametric tests, we find that the number of signals acquired in situation  $2B(2 : 0)$  is less than in situation  $0B(3 : 1)$ . This is true for both the BSs data ( $U = 1.680$ ,  $p = 0.046$ ) and the WSs data ( $z = 1.542$ ,  $p = 0.062$ ), although the results are only marginally significant with the WSs data.*

*The reduced-form results obtained from our regressions are only marginally significant. This is true when estimating a random effects Poisson model (BSs:  $p = 0.065$ , WSs:  $p = 0.051$ ) or a random effects Tobit model (BSs:  $p = 0.058$ , WSs:  $p = 0.001$ ).<sup>27</sup> The top panel of figure 3 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 our two subsamples.*

As discussed previously, exogenous variation in the realizations of the free signals that subjects observe before the information acquisition decision allow us to sample the relatively rare knife-edge case in which the updated prior implies indifference. Specifically, in situations  $4B$  and  $2B(1 : 1)$  subjects are promised free information in the future but have a prior of  $\beta_\omega = 1/2$  when purchasing signals. In situation  $0B(2 : 2)$  (where the prior is also  $\beta_\omega = 1/2$ ), no additional free signals are forthcoming. Our model predicts that the promise of free signals in the future will reduce costly information acquisition in this case. This prediction is borne out in our data.

**Result 2.1 (extensive margin):** *When  $\beta_\omega = 1/2$ , subjects are less likely to purchase signals when they know that they will observe free signals in the future. That is, the probability of purchasing signals is higher in  $4B$  or  $2B(1 : 1)$  than in  $0B(2 : 2)$ .*

**Support:** *We reject the null hypothesis that the probability of acquiring any information is equal across the three relevant situations ( $\chi^2 = 13.862$ ,  $p < 0.001$  for BSs data; Friedman = 195.2,  $p < 0.0001$ ; for WSs data). When restricting attention to pairwise comparisons, we reject the null that  $4B \geq 0B(2 : 2)$  (BSs data:  $U = -3.813$ ,  $p < 0.0001$ ; WSs data:  $z = -5.605$ ,  $p < 0.0001$ ), that  $2B(1 : 1) \geq 0B(2 : 2)$  (BSs data:*

<sup>27</sup>Appendix table A.4 presents the full results for the regression estimates.

$U = -1.680, p = 0.046$ ; WSs data:  $z = -3.953, p < 0.0001$ ), and that  $4B \geq 2B(1 : 1)$  (BSs data:  $U = -2.2253, p = 0.013$ ; WSs data:  $z = -3.341, p < 0.001$ ).

*Our reduced-form models, reported in table A.3 and illustrated in the bottom panel of figure 2, also strongly support these results.*<sup>28</sup>

This result on the extensive margin is in line with the predictions of our model. However, it could be the case that the promise of free balls lowers the probability that a subject purchases signals while still increasing the quantity of signals purchased. As such, we turn to the intensive margin.

**Result 2.2 (intensive margin):** *When  $\beta_\omega = 1/2$ , the promise of free signals to be observed after the information acquisition decision leads subjects to purchase fewer signals at cost. That is, fewer signals are purchased in  $4B$  or  $2B(1 : 1)$  than in  $0B(2 : 2)$ .*

**Support:** *We reject the null hypothesis of equal information acquisition across the three relevant situations. Using the BSs data, this result is marginally significant ( $\chi^2 = 4.752, p = 0.093$ ), while it is highly significant (Friedman = 194.96,  $p < 0.0001$ ) when using the WSs data. When comparing situations with the largest difference in the number of free signals promised in the future ( $4B$  versus  $0B(2 : 2)$ ), the results are highly significant. We reject the null that  $4B \geq 0B(2 : 2)$  (BSs data:  $U = -2.260, p = 0.0119$ ; WSs data:  $-2.870, p = 0.002$ ). When there are only two additional signals promised in the future, the results have the predicted sign, but the significance of the nonparametric tests are inconsistent between the BSs and the WSs data. We can only reject the null that  $2B(1 : 1) \geq 0B(2 : 2)$  using the WSs data (BSs data:  $U = -1.194, p = 0.116$ ; WSs data:  $z = -0.898, p = 0.019$ ). Similarly, for the null that  $4B \geq 2B(1 : 1)$ , the results are only significant using the WSs data (BSs data:  $U = -0.909, p = 0.182$ ; WSs data:  $z = -2.431, p < 0.007$ ). This is driven by the higher level of statistical power available in the WSs data.*

*Our reduced-form results are consistent with the results of the nonparametric tests. Plots of the relevant coefficients are contained in the bottom panel of figure 3.*<sup>29</sup> *Note that the regression analysis consistently demonstrates that more signals are purchased in situation  $4B$  than in situation  $0B(2 : 2)$ . This is true for estimates obtained using a random effects Poisson model or a random effects Tobit model. However, when the comparison is less stark, we only reject the other null hypotheses ( $2B(1 : 1) \geq 0B(2 : 2)$  and  $4B \geq 2B(1 : 1)$ ) using the WSs data.*

In addition to the main hypotheses, our data support the prediction that promising free information in the future will not increase decision quality when priors are asymmetric. Despite differences in the amount of information acquired, the proportion of time the subjects choose the correct jar is no higher in  $2B(2 : 0)$  than in  $0B(3 : 1)$ .<sup>30</sup> However, in the 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

<sup>28</sup>The only exception is  $2B(1 : 1) \geq 0B(2 : 2)$ , where the difference is not statistically significant with the BSs data; however, it is highly significant ( $p < 0.001$ ) with the WSs data.

<sup>29</sup>Table A.4 presents the full results for the regression estimates.

<sup>30</sup>Nonparametric tests:  $p = 0.368$  for BSs,  $p = 0.6146$  for WSs. Reduced-form regressions:  $p > 0.307$  for BSs and  $p > 0.104$  for WSs.

guesses. We find support for this when comparing the most extreme situation with symmetric priors. That is, predictive accuracy is higher when four signals are promised relative to the case when no additional free signals are forthcoming ( $0B(2:2)$  versus  $4B$ ).<sup>31</sup>

As demonstrated above, we find broad support for the theoretical predictions of our model. Most importantly, we find empirical support for the theoretically predicted ranking of information acquisition in the six situations we consider. However, our model fails to predict the observed level of information acquisition; our subjects consistently overacquire information. Such overacquisition is commonly observed in experimental settings, and our results are in line with those of Page and Siemroth (2017), who employ a similar environment to study information acquisition in asset markets. Similar overinvestment in information has been observed by Gretschko and Rajko (2015) in an auction context and by Chen and He (2021) and Hakimov et al. (2021), who study information acquisition in matching.

There is little consensus in the literature on the reasons for this rather consistent observation, and a number of possible explanations have been proposed. For example, Bhattacharya et al. (2017) and Meyer and Rentschler (2022) observe overinvestment in information in voting environments and consider quantal response equilibrium as a possible explanation. This, however, cannot explain our data as we focus on a decision-theoretic environment. Gretschko and Rajko (2015) hypothesize that overacquisition of information could be driven by regret avoidance.<sup>32</sup> Chen and He (2021) consider several explanations and empirically assess the drivers of willingness to pay for information in a matching contest. They find that subjects' willingness to pay increases in their beliefs about the willingness to pay of others and that curiosity also plays a significant role. Page and Siemroth (2017) find that information acquisition increases in subject endowments and decreases in risk aversion and experience in financial markets. They also find that observing inconclusive information leads to more information acquisition.

Conclusively determining the ability of most of these explanations to explain the overacquisition observed in our data would require additional research that would systematically vary aspects of the environment. This is beyond the scope of this paper, although we view it as a promising avenue for future research. However, since the second task of our experiment elicited the risk preferences of subjects, we can determine whether observed behavior can be explained by risk aversion. Unfortunately, the answer is negative. The difference between the predicted level of information acquisition of a risk-neutral agent and of an agent exhibiting the median level of risk aversion observed in our experiment is economically trivial.<sup>33</sup> Further, accounting for this median level of risk aversion often reduces the value of information. This demonstrates that risk aversion cannot explain our data on aggregate. Below, we will also demonstrate that observed risk aversion cannot explain the observed patterns in our data when accounting for individual heterogeneity in responses.

---

<sup>31</sup>BSs:  $U = 2.136$ ,  $p = 0.016$ , WSs:  $z = 3.607$ ,  $p < 0.001$ . We also estimate the probability of correctly guessing the state 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. The results are presented in table A.5. Again, we find consistent support for the predictions.

<sup>32</sup>Interestingly, regret avoidance has also been proposed as an explanation for information avoidance; see, for example, Golman et al. (2017).

<sup>33</sup>The largest change in willingness to pay is for the first signal in situation  $2B(2 : 0)$ . Here the willingness to pay of the risk-averse agent is 1.72 less than in the risk-neutral case. All other differences in willingness to pay are less than 1 in magnitude. We assumed a CRRA utility function with the risk-aversion coefficient, consistent with the median choice observed in our data for this analysis.

## 5.1 Individual Behavior

In this subsection we explore the heterogeneity of individual-level behavior. We begin with a descriptive analysis of individual heterogeneity and then focus on the WSs data to analyze individual-level behavior across treatments.

We start by checking whether individual subjects respond to experimental parameters in a manner consistent with our theoretical predictions. In the WSs data only 2/84 (2.38 percent) subjects never acquire any information in all 48 periods; 2/84 (2.38 percent) subjects buy at least one signal in all 48 periods. The median [10th, 90th percentile] individual buys at least one signal in 42.7 percent [10.4 percent, 81.3 percent] of the periods.<sup>34</sup> Most subjects also seem to have a downward-sloping demand for signals: for 70/84 (83.3 percent) of our subjects, we estimate a statistically significant negative effect of the cost on information acquisition (see figure B.5).<sup>35</sup>

In a more subtle test of the theory, most subjects seem to understand the nonmonotonic value of signals, in which purchasing an even number of signals is never predicted. In fact, conditional on purchasing a positive number of signals, 11/84 (13.1 percent) of subjects always acquire an odd number of signals, whereas no subject chooses to always acquire an even number of signals. Again, conditioning on purchasing a positive number of signals, the median [10th, 90th percentile] subject acquires an odd number of signals 73.5 percent [52.1 percent, 100 percent] of the time. The number of subjects who, when purchasing, always purchase an odd number of signals more than double to 23/84 (27.4 percent) when excluding the first 12 periods, suggesting possible learning with experience.<sup>36</sup>

For the rest of this subsection, we rely on the WSs data to examine heterogeneity in treatment effects across individuals. We focus on the case of the asymmetric prior where  $\beta_\omega = 9/13$ , and theory predicts that subjects will acquire more information if free information is promised in the future. For each individual  $i$ , we compare the mean number of signals purchased across all periods in which subject  $i$  was in situation  $2B(2 : 0)$  to the mean number of signals purchased when subject  $i$  was in situation  $0B(3 : 1)$ .<sup>37</sup> We separately consider the comparison between the two knife-edge situations with a symmetric prior of  $\beta_\omega = 1/2$ :  $2B(1 : 1)$  and  $0B(2 : 2)$ .

To provide a meaningful comparison, for each subject, we compare the observed individual-level treatment effect to the predicted individual-level treatment effect, conditioning on the realized cost of information subjects actually faced in each period. We do this to account for the fact that subjects faced different costs across periods since the cost of signals is randomized in each period. For each subject, we directly compare

---

<sup>34</sup>Among the participants in the first wave (UFM subject pool), 1/72 (1.4 percent) never purchases any information and 8/72 (11.11 percent) always acquire at least one signal. The median [10th, 90th percentile] individual buys at least one signal in 66.7 percent [25 percent, 100 percent] of periods.

<sup>35</sup>When restricting this analysis to subjects in the first wave (UFM subject pool), we see similar results. Since we have half the observations for this subsample (24 instead of 48), we use 90 percent, rather than 95 percent confidence intervals, to claim statistical significance. Figure B.6 illustrates that only 7/72 (9.7 percent) subjects do not have downward-sloping demand.

<sup>36</sup>Among the subjects in the first wave (UFM subject pool), 13/72 (18.1 percent) always purchase an odd number of signals; 28/72 (38.9 percent) fully comply with the heuristic when restricting to the second half. Only 1/72 always purchase an even number of signals. The median [10th, 90th percentile] subject, when purchasing signals, acquires an odd number of signals 78.6 percent [50 percent, 100 percent] of the time.

<sup>37</sup>In a given treatment, the number of periods in which a subject would be in each of these situations varies and depends on the realizations of the free signals.

the observed treatment effect to the predicted treatment effect, conditioning on the realized cost of information.

On average, pooling across all subjects, we predict a positive mean difference of 0.141: more information is predicted to be acquired when future free information is promised. However, there is considerable heterogeneity at the individual level: predicted difference in means is positive for 46 (54.8 percent) subjects, 0 for 17 (20.2 percent), and negative for 21 (25 percent). (To reiterate, this happens because some subjects faced lower costs in  $0B(3 : 1)$  than in  $2B(2 : 0)$ .)

Figure 4 illustrates (for each of the 84 subjects from the wave 2 sessions) the observed difference in individual-level information acquisition when  $\beta_\omega = 9/13$ . Solid circles indicate the average number of signals purchased in situation  $2B(2 : 0)$  less the average number of signals purchased in situation  $0B(3 : 1)$ . Empty diamonds indicate the respective predicted difference. For comparison, figure 5 presents the same for the knife-edge case of  $\beta_\omega = 1/2$ . The white  $x$ s (+s) denote the two subjects who never (always) acquire any signals.

In both of the preceding figures, the coloring of the solid circles denotes the relationship between the observed treatment effect and the predicted treatment effect. Black circles denote individuals whose observed difference in mean behavior across treatments has (weakly) the same sign as the predicted difference. Red circles denote individuals where the observed and predicted differences have the opposite sign. The vertical lines connecting the circle and diamonds indicate the distance between the observed and predicted differences. Red lines denote instances in which the observed difference in mean behavior across treatments is less than the predicted difference. For subjects whose observed difference has the same sign as the predicted difference (black solid dots), the red line indicates that they *underresponded* to the theoretically predicted treatment difference. Gray lines indicate instances in which the observed difference has the same sign as the predicted difference but the magnitude is larger than predicted.

Of note is the fact that when  $\beta_\omega = 9/13$ , the individual-level treatment effect fails to match the sign of the predicted effect for only 17/84 of subjects. Further, when  $\beta_\omega = 1/2$ , only 27/84 subjects fail to match the sign of the observed and the predicted difference. That is to say, while there is significant heterogeneity in individual-level treatment effects, our results are not driven by a small number of outliers. Rather, the behavior of the majority of subjects are qualitatively in line with the theoretical predictions.

We next explore how well the individual-level treatment effects correlate with the predicted individual-level treatment effects. Figure 6 presents a scatter plot and a linear fit (with the associated 95 percent confidence interval) of predicted individual-level treatment differences and the predicted individual-level treatment differences. The left panel illustrates situations where  $\beta_\omega = 9/13$ , and the right panel illustrates situations in which  $\beta_\omega = 1/2$ . There is a strong, positive correlation both when  $\beta_\omega = 9/13$  ( $\rho = 0.517, p < 0.001$ ) and when  $\beta_\omega = 1/2$  ( $\rho = 0.688, p < 0.001$ ).

These results are further backed by the regression analysis reported in tables A.6 and A.7. The tables report ordinary least squares (OLS) estimates of the individual-level differences in behavior across treatments for  $\beta_\omega = 9/13 : 2B(2 : 0) - 0B(3 : 1)$  and  $\beta_\omega = 1/2 : 2B(1 : 1) - 0B(2 : 2)$ , respectively. In both tables, the dependent variable in columns 1–2 (4–5) is individual differences in mean (median) information acquisition across situations. The dependent variable in column 3 (6) is the individual observed deviation

from predictions. Column 2 (4) shows that the predicted difference in means (medians) is positive and statistically significantly correlated with the observed difference. The difference in means, however, is smaller than (and statistically significantly different from) one. For the difference in medians (column 4), we cannot reject the null hypothesis that the coefficient is equal to one.

Regarding individual-level characteristics that may correlate with treatment effects, we observe that the only robust and consistent difference is related to the measured risk preferences, as measured by the number of safe options the subject chose in the risk elicitation task.<sup>38</sup> In particular, we find that treatment differences are smaller for those who are more risk averse. However, it is important to note that the magnitude of these coefficients is small.

The main takeaway is that we observe a strong and positive correlation between individual-level observed and predicted difference in means (medians). Thus, despite individual heterogeneity, we find that a large majority of subjects exhibit behavior that differs across treatments in the same direction as the theoretical predictions and that individual differences in behavior across treatments are strongly correlated with the theoretical predictions.

## 6 Discussion and Further Research

We present results of a laboratory experiment on the impact a promise of future information may exert on individual information acquisition effort. We observe that offering future information encourages greater costly information acquisition. Furthermore, when we explore the knife-edge case of symmetric priors, we observe (as predicted) that the promise of future information reverses this result. At the aggregate level, the differences across treatments are more pronounced on the extensive than on the intensive margin: a 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.

The aggregate results are not driven by a few subjects. When focusing on individual-level treatment differences using the within-subject variation, we find that behavioral support for the theory is widespread. Out of the 80 subjects who exhibit responsiveness to experimental parameters in their information acquisition decisions, about 65 percent show a difference in information acquisition across the laboratory environments going in the same direction as the theoretical prediction. Furthermore, we find a strong correlation between the observed and predicted differences in mean information acquisition decisions across treatments. Thus, we observe strong evidence to support the theory, both at the aggregate and the individual level. 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.

Though a promise of free information may create incentives for costly information acquisition that would not exist otherwise, it does not improve the quality of individual decisions, nor the welfare of the

---

<sup>38</sup>The number of safe options is the number of times a subject chose the relatively safe option A, rather than the relatively risky option B, in the 10 choices from the risk elicitation task. Since subjects can switch back and forth between option A and option B (indicating inconsistent preferences, confusion, or indifference), we also control for whether subjects switched multiple times. In addition, we control for whether a subject switched at all.



decisionmaker. This confirms the theoretical prediction that follows straightforwardly from Blackwell (1962).<sup>39</sup> However, in group settings, such as juries or committees, if we can use the promise of future information to induce additional information acquisition effort by individuals, we may expect it 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 a trial until all evidence is presented. In Hannaford et al. (2000), the authors show intriguing evidence of increased drop out by jurors in a field experiment in which this prohibition was relaxed, lending support for the hypothesis we have just laid out. Consequently, in terms of institutional design, delay in information provision may be sufficient to discourage “informational dropout,” situations in which agents choose to forgo attentional effort and make decisions based entirely on their prior beliefs. It may also be a useful tool in experimental design as it could be used to avoid excessive dropout by subjects that has been observed in some previous experimental studies (Elbittar et al., 2020).

Finally, this study poses some questions to the important literature that evaluates the impact of free information provision and suggests that a more rigorous framework is needed. Since the information provided in most studies cannot be anticipated at the moment of making information acquisition decisions, it would seem that the impact of information provision may sometimes be ambiguous. Thus, a good applied information provision design should take into account the nonconcavity in the value of information, make explicit the underlying assumptions behind the baseline priors, and incorporate the impact of established biases (such as, e.g., base rate neglect). Our paper aims to provide a laboratory contribution to answering at least some of these questions.

---

<sup>39</sup>This is striking and an important aspect of our analysis since policymakers often seek to induce individual informedness under the presumption that this is a social good.

## 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.
- 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 Acemoglu, 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.
- Chen, Yan and Yinghua He, "Information acquisition and provision in school choice: an experimental study," *Journal of Economic Theory*, 2021, p. 105345.
- 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*, 2020, 174, 332–359.
- Feltoich, Nick, "Nonparametric tests of differences in medians: comparison of the Wilcoxon–Mann–Whitney and robust rank-order tests," *Experimental Economics*, 2003, 6 (3), 273–297.

- 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.
- Golman, Russell, David Hagmann, and George Loewenstein, “Information avoidance,” *Journal of Economic Literature*, 2017, 55 (1), 96–135.
- Gretschko, Vitali and Alexander Rajko, “Excess information acquisition in auctions,” *Experimental Economics*, 2015, 18 (3), 335–355.
- Grosser, Jens and Michael Seebauer, “The curse of uninformed voting: An experimental study,” *Games and Economic Behavior*, 2016, 97, 205–226.
- 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.
- Hakimov, Rustamdjan, Dorothea Kübler, and Siqi Pan, “Costly information acquisition in centralized matching markets,” 2021.
- 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.
- Kanazawa, Satoshi, “What do we do with the WEIRD problem?,” *Evolutionary Behavioral Sciences*, 2020, 14 (4), 342.
- 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.
- Meyer, Jacob and Lucas Rentschler, “Abstention and informedness in nonpartisan elections,” *Working Paper*, 2022.
- Page, Lionel and Christoph Siemroth, “An experimental analysis of information acquisition in prediction markets,” *Games and Economic Behavior*, 2017, 101, 354–378.
- 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.











































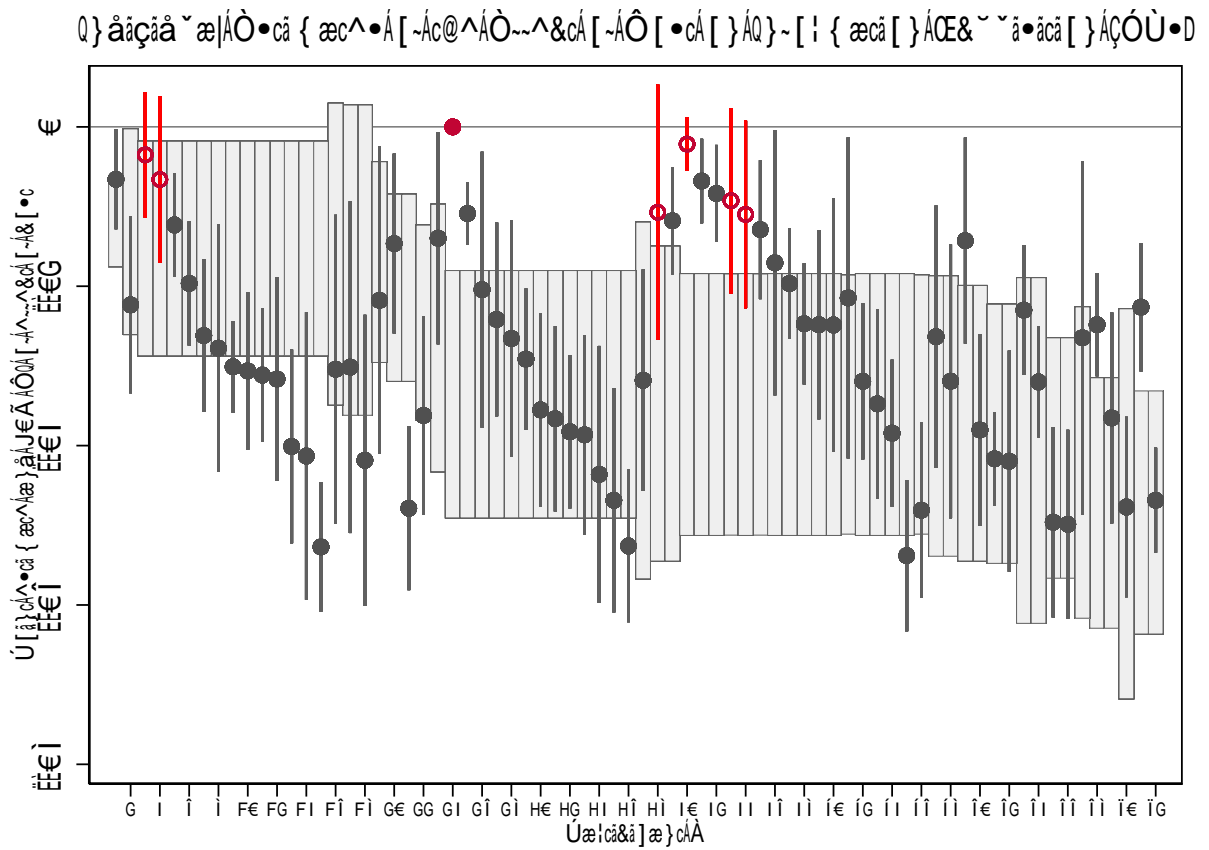


Figure B.6. Individual Estimates of the Effect of Cost on Information Acquisition Decisions Using between-Subject Data from Wave 1 (UFM Sample). The figure plots point estimates with 90 percent confidence intervals. The shaded area corresponds to 90 percent confidence intervals for the predicted effect of cost on information acquisition under risk neutrality.

## C Instructions

This section presents the script used to run the wave 2 sessions and the slides to present instructions.<sup>40</sup> 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 section D.2 presents the control questions that followed the instructions. Appendix section D.3 presents the script used to conduct wave 2 sessions.

### C.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 your 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.

---

<sup>40</sup>We choose instructions for wave 2 since they are in English. Instructions for wave 1 are available from the authors upon request.



## **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).

Next, we will explain this part of the experiment in more 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:

$$E\$1,000 - \text{COST} \cdot (\# \text{ of balls purchased})$$

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

$$E\$300 - \text{COST} \cdot (\# \text{ of balls purchased})$$

## 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:

$$E\$1,000 - \text{COST} \cdot (\# \text{ of balls purchased})$$

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

$$E\$300 - \text{COST} \cdot (\# \text{ of balls purchased})$$

## C.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.