Expectation Formation under Uninformative Signals*

Pascal Kieren† and Martin Weber‡

This draft: October, 2020

Abstract

How do individuals process non-diagnostic information? The neoclassical theory of probabilistic beliefs assumes that people update their prior beliefs according to Bayes' Theorem as new (relevant) information arrives. This paper provides experimental evidence that individuals update their prior beliefs even after observing uninformative signals. Importantly, the direction in which they update depends on the valence of the signal. Prior beliefs become more optimistic after desirable uninformative signals and more pessimistic after undesirable uninformative signals. This bias in information processing becomes even more severe when the valence of a signal is at odds with individuals' prior beliefs about the desirability and in situations in which individuals are intrinsically motivated to hold a particular belief. Our results provide novel insights why individuals form and entertain false beliefs in environments where potentially new information is easily accessible but costly to verify (e.g. online media).

Keywords: *Belief Formation, Information Processing, Uninformative Signals, Bayes' Theorem*

JEL Classification: *C91, D12, D81, D83, D91*

---

* The experiment in this paper was pre-registered in the Credibility Lab of the Wharton University of Pennsylvania, accessible under https://aspredicted.org/.

† Pascal Kieren, University of Mannheim, Department of Finance, E-mail: pascal.kieren@gess.uni-mannheim.de, corresponding author.

‡ Martin Weber, University of Mannheim, Department of Finance and the Centre for Economic Policy Research, E-mail: weber@bank.bwl.uni-mannheim.de

1 Introduction

Probabilistic judgements are a central feature of any theory that involves decision-making under risk. As such, errors in probabilistic reasoning matter for essentially any economic decision that involves risk, including retirement, investments, purchasing insurance, or attaining various degrees of education. In the textbook model of Bayesian Updating, individuals update their prior beliefs according to Bayes' Rule upon receipt of new information. In this model, signals which do not carry relevant information about the objective state of the world play no role and are treated as if no signal occurred.

In reality however, many information structures are complex, generating signals that are often noisy and difficult to ascribe to one particular state of the world. Additionally, new information is rarely processed as being purely informative. Instead, individuals frequently have preferences over which state of the world is true, effectively generating an interaction between beliefs and preferences (Eil & Rao, 2011; Möbius et al., 2014). This interaction may lead to environments, in which information signals carry no information about an underlying state of the world, but which nonetheless appear either desirable or undesirable. While Bayes' Theorem would prescribe that individuals do not update their prior beliefs in response to such *uninformative* signals, it is unclear whether individuals can correctly discern belief-relevant information from their preferences.

Taking this observation as a point of departure, we conduct an experimental study, in which we investigate how agents process signals which are non-diagnostic about the objective state of the world but which are either desirable or undesirable in the payoffs they generate. In the experiment which partly builds on Grether (1980), subjects have to incorporate a series of information signals into their beliefs to forecast the distribution of a risky asset. The risky asset can generate three outcomes from one of two distributions, a bad distribution and a good distribution. The outcomes can be ranked according to their associated payoff (high, medium, and low). In the good distribution, the high outcome occurs with the highest probability, while the low outcome occurs with the lowest probability. In the bad distribution, probabilities of the high and low outcome are reversed. Following this logic, the high outcome signals that the good distribution is more likely, whereas the low outcome signals that the bad distribution is more likely. Importantly, the medium outcome always occurs with the same probability independent of the underlying distribution. In other words, the medium outcome provides no opportunity to learn about the true state of the risky asset and will subsequently be referred to as an *uninformative* signal.

Over the course of ten rounds, subjects observe random draws from one of the two distributions and have to make a probability forecast about the likelihood that the asset is drawing from the good distribution. In our experiment, we have two key treatment variations which we exogenously vary in a between-subject design. The first treatment variation allows us to investigate how the valence of uninformative signals affects individuals' updating behavior. In the positive treatment, the uninformative signal pays a positive payoff, whereas in the negative treatment, the uninformative signal pays a negative payoff. The second treatment variation concerns the motivation to provide correct forecasts. In the passive treatment, subjects are asked to forecast the distribution of the risky asset after each draw and are thus only motived to be accurate in their probability forecasts. In the active task, subjects additionally decide each round between investing in the risky asset or a riskless security which always pays the intermediate outcome. In this condition, subjects are motivated to be accurate in their forecasts and to maximize their payoffs.

In our experimental setting, we have direct control over objective expectations and can compare them to participants' subjective beliefs. Importantly, the distributions from which information is drawn are constructed in a way, that the medium outcome does not provide any information about whether subjects are currently drawing from the good or the bad distribution. As such, a Bayesian agent in our setting should not update his prior beliefs after observing an uninformative signal, independent of whether the signal is in the positive domain or in the negative domain. This allows us to disentangle the valence from the informational content of a signal and to document systematic errors in the belief formation process.

We find that individuals strongly and systematically update their prior beliefs after observing signals that are uninformative of the objective state of the world. In contrast to Bayesian behavior, individuals fail to fully extract belief-relevant information. Whereas they update their priors in on average by about 7.45 percentage points after observing informative signals, they also update their priors by 2.21 percentage points after observing uninformative signals. In relative terms, individuals adjust their priors with about 30 % of the strength as if the observed signal would carry information.

Second, we find that the direction in which individuals update their beliefs strongly depends on the valence of the observed signal. In particular, individuals tend to form more optimistic beliefs about the objective state of the world after observing positive uninformative signals, whereas they form more pessimistic beliefs after observing negative uninformative signals. This effect becomes even more pronounced when individuals observe uninformative signals in an environment in which their beliefs matter for a payoff-relevant decision. To the best of our knowledge, our study is the first test of whether individuals can distinguish their preferences from belief-relevant information in their belief formation. Additionally, we show that the effect is not driven by a few individuals who overreact to the valence of uninformative signals, but rather a general phenomenon. After observing informative signals subjects only occasionally make updating mistakes that are directionally inconsistent with Bayes' Rule (e.g. becoming more optimistic after a bad signal). However, after observing uninformative signals, subjects wrongly update their beliefs in about 68 % of the cases.

Third, as underlying mechanism we identify that individuals tend to process noisy information signals in a reference-dependent manner dictated by their prior beliefs. They fail to correctly identify that uninformative signals do not carry information about the objective state of the world and update their beliefs based on the valence of the signal relative to their current prior expectations. In particular, subjects who hold optimistic prior beliefs about the state of the risky asset (i.e. subjects who belief the good outcome is more likely to occur) only weakly increase their beliefs when the uninformative signal is positive (but in magnitude smaller than the good signal), but strongly decrease their beliefs when the signal is negative. Similarly, subjects who hold pessimistic prior beliefs about the state of the risky asset only weakly increase their beliefs when the uninformative signal is negative (but in magnitude smaller than the bad signal), but strongly increase their beliefs when the signal is positive.

Research on errors in probabilistic reasoning has a long-standing tradition (Philips & Edwards, 1966; Tversky & Kahneman, 1971, 1974). Implications of biased reasoning following new information have been studied in diverse contexts such as in psychologists' interpretation of diagnostic tests (Meehl and Rosen; 1955), doctors' diagnoses of patients (Eddy, 1982), courts' judgments in trials (Tribe, 1971), or ideological conflicts and political discussions (Kahan, 2013). This article contributes to the literature by identifying an error in how individuals process information signals which provide no relevant learning opportunity about an objective state of the world, but which are nonetheless desirable or undesirable in the payoffs they generate. Our findings most closely relate to earlier studies which investigate base-rate neglect in response to uninformative descriptions of personality sketches (see e.g. Kahneman & Tversky, 1973; Wells & Harvey, 1978; Ginosar & Trope, 1980, 1987; Fischhoff & Bar-Hillel, 1984). These experiments typically consist of instructions which are framed to be irrelevant for judging the likelihood that a person belongs to a particular job group and find that individuals by and large draw inferences from such descriptions. Whereas these studies also examine how uninformative descriptions affect individuals' judgement about underlying probabilities, they are fundamentally different from ours as they do not investigate the influence of uninformative signals in dynamic belief updating problems. Perhaps closest to our study is the study by Troutman and Shanteau (1977), who investigate the effect of different non-diagnostic samples in bookbag-and-poker-chip experiments. They find that non-diagnostic samples result in less extreme probability assessments, as individuals effectively average across all observed signals, thereby mixing both informative and uninformative signals. Yet, different from existing work, our study emphasizes the critical role of preferences in the processing of uninformative signals. We show that, depending on the valence of the signal and individuals' prior beliefs, non-diagnostic signals can also lead to *more extreme* responses. As such, uninformative signals can not only lead to systematically biased beliefs whenever desired or undesired outcomes are non-indicative of the true state of the world, but also reinforce wrongly entertained beliefs.

Our paper also relates and contributes to the literature on prior-biased inference especially in the context of confirmation bias (e.g. Charness & Dave, 2017) and belief-polarization (e.g. Lord et al., 1979; Kahan, 2013; or Benoît & Dubra, 2018). This literature finds that individuals have a tendency to seek, interpret, and use evidence in a manner biased towards current beliefs. In belief polarization experiments, subjects with different priors are typically presented with the same mixed signals, causing their beliefs to move further apart. In these studies, signals are usually informative although noisy, effectively giving room for different interpretations. Our results highlight that even in settings in which signals are non-diagnostic of an objective state of the world, beliefs might drift apart if individuals have different priors and assign a different level of valence to the signal. In the presence of an increasing number of information sources, the mechanism presented here might further reinforce polarized beliefs.

Finally, our paper also contributes to the broad literature on motivated beliefs, which argues that beliefs are adjusted differently depending on the valence of the observed signal. Especially in the context of self-relevant beliefs, individuals appear to asymmetrically process self-serving information, putting more weight on positive than on negative information (see e.g. Eil & Rao, 2011; Sharot et al., 2011; or Zimmerman, 2020). While the beliefs we elicit in our study are not self-relevant, they are nonetheless motivated as participants are motivated to believe that the risky asset is in the good state, because the good state is more likely to result in greater payoffs. For informative signals, we find that individuals update their beliefs regarding the state of the risky asset to a greater extent following positive information than negative information. However, similar conclusions cannot be drawn regarding uninformative signals. Here, individuals appear to process the signals in a rather symmetric manner for priors close to 50-50, becoming more optimistic after positive uninformative signals and more pessimistic after negative uninformative signals. Yet, once individuals become pessimistic, they also start to asymmetrically update their beliefs, strongly overreacting to positive uninformative signals and mostly neglecting negative uninformative signals. As in the model proposed by Bénabou (2013), this mechanism might suggest that individuals want to quickly revert very pessimistic priors to preserve anticipatory utility from putting a higher probability on the good state.

The remainder of the paper is structured as follows. Section 2 offers a stylized formal framework that motivates the experimental design and structures the empirical analysis. Section 3 presents evidence that subjects update their prior beliefs even after observing uninformative signals based on the valence of the signal and explores potential mechanisms underlying this phenomenon. Section 4 concludes.

2 Conceptual Framework

A. Setup

This section presents a stylized framework to guide the design of the experiment and to structure the main part of the empirical analysis. The underlying mechanics of the framework directly build on a reduced-form model originally introduced by Grether (1980). To keep the focus on the processing of uninformative signals, we assume only two possible states of the world, a good state (denoted as G) and a bad state (denoted as B). Consider a decision-maker (DM) who wants to learn about the current state of the world. The agent's prior beliefs are denoted by $p(g)$ and $p(b)$. To decide which state of the world is more likely, the DM receives a number of signals S, in which each signal s_t can either be informative of a good state (signal g) or of a bad state (signal b). Additionally, the DM may also receive uninformative signals (signal u), which are neither indicative of a good state nor of a bad state. As the DM observes a new signal, she updates her prior beliefs according to the following function:

$$\pi(G|S) = \frac{p(S|G)^c p(G)^d}{p(S|G)^c p(G)^d + p(S|B)^c p(B)^d} \qquad (1.1)$$

$$\pi(B|S) = \frac{p(S|B)^c p(B)^d}{p(S|G)^c p(G)^d + p(S|B)^c p(B)^d} \qquad (1.2)$$

where $p(.)$ refers to a true conditional probability, $\pi(.)$ refers to an agent's (potentially biased) belief, and $c, d \geq 0$. The parameter c governs the (biased) use of likelihoods, while the parameter d governs the (biased) use of prior beliefs. To interpret the magnitudes of c and d, we follow Benjamin (2019), and write the model in the posterior-odds form, dividing Eq. (1.1) by Eq. (1.2):

$$\frac{\pi(G|S)}{\pi(B|S)} = \left[\frac{p(S|G)}{p(S|B)}\right]^c \left[\frac{p(G)}{p(B)}\right]^d.$$

In this equation, $c < 1$ corresponds to updating as if the signals provided less information about the state (underinference), while $c > 1$ corresponds to updating as if the signals provided more informative than they do (overinference). Similarly, $d < 1$ corresponds to treating the priors as less informative than they are (also referred to as base-rate neglect), while $d > 1$ corresponds to the opposite. The model nests Bayes' Theorem as a special case, in which $c = d = 1$.

To infer the underlying state of the world, consider that a DM receives each period t a new signal, which can either be good, bad, or uninformative. In a signal structure where only two signals carry information about the underlying state of the world, the conditional probability of being in the good state given the signal history at time t ($\pi(G|S)$) can be calculated as follows:

$$\pi^{Bayes}(G|S) = \frac{\theta^{z_t}}{\theta^{z_t} + (1-\theta)^{z_t}}, \qquad z_t = g_t - b_t$$

where g_t (b_t) denotes the number of good (bad) signals that have been observed until period t and z_t is the difference between both. The parameter $\theta \in [0,1]$ captures the diagnosticity of an informative signal. Since a Bayesian DM would neglect uninformative signals, only the difference of good and bad signals is of relevance. Additionally, note that he is indifferent regarding the order in which the signals occur.

Following Charness and Dave (2017), we make use of the fact that the natural log of the odds ratio within a round ($\ln\left(\frac{\pi^{Bayes}(G|z_t)}{\pi^{Bayes}(B|z_t)}\right)$) for a Bayesian is given by[4]:

$$\pi_t^{Bayes} = \ln\left(\frac{\pi(G|s_1, \dots, s_t)}{\pi(B|s_1, \dots, s_t)}\right) = \ln\left(\frac{\theta}{1-\theta}\right) \cdot z_t$$

[4] A detailed explanation and derivation is provided in Appendix C.

As such, the Bayesian log-odds ratio is updated by $\pm\theta \cdot z_t$ after each new signal and – in contrast to Bayes probability – linear in the number of signals. To make the interpretation easier, we take the first-difference of both sides of the equation, yielding:

$$\Delta\pi_t = \pi_t - \pi_{t-1} = \ln\left(\frac{\theta}{1-\theta}\right) \cdot \Delta z_t,$$

where $\Delta\pi_t \in \{-\theta, \theta\}$ and $\Delta z_t \in \{-1,0,1\}$. The interpretation of $\Delta\pi_t$ is straightforward. If the DM observes a new good (bad) signal, then the Bayesian log-odds ratio is updated by θ $(-\theta)$. If the DM observes an uninformative signal (i.e. $\Delta z_t = 0$), then the Bayesian log-odds ratio remains constant.

To incorporate that a non-Bayesian DM may falsely incorporate an uninformative signal in his belief updating process, we consider the possibility that c may not only depend on the information of a signal, but also on the valance. In our setting, valance can be broadly defined as any signal that does not help the DM to learn about the current state of the world but which provides either positive or negative utility (e.g. through a payoff or other factors that might be relevant for the DM):

$$\frac{\pi(G|S)}{\pi(B|S)} = \left[\frac{p(S|G)}{p(S|B)}\right]^{c_0 + I\{u|\ desirable\} \cdot c_1 + I\{u|\ undesirable\} \cdot c_2} \left[\frac{p(G)}{p(B)}\right]^d, \quad (1.3)$$

where $I\{u|\ desirable\}$ equals 1 if $s = u$ and the signal is perceived as desirable and $I\{u \mid undesirable\}$ equals 1 if $s = u$ and the signal is perceived as undesirable. In this case, we obtain three reduced-form parameters which describe biased inference: c_0 captures inference of informative signals, while c_1 and c_2 capture uninformative signals which are desirable or undesirable, respectively.

Equation (1.3) will be the core expression to investigate individuals' propensity to update after uninformative signals. It nests the Bayesian prediction that priors are fully incorporated in the belief formation process (i.e. $d = 1$) and that individuals respond with a coefficient of $c_0 = 1$ to informative signals. Finally, since a Bayesian DM would not update his prior beliefs after an uninformative signal (i.e. $\Delta z_t = 0$), we would expect that $c_1 = c_2 = 0$.

B. Experimental Design

To study the degree to which individuals' belief formation process is affected uninformative signals, we require an environment in which (i) individuals repeatedly incorporate signals with varying degrees of information into their beliefs; (ii) Bayesian beliefs can be clearly identified; (iii) treatment variations allow the exogenous variation of the desirability of uninformative signals; (iv) holding positive/negative beliefs has a value in and of itself; and (v) the belief elicitation is incentive-compatible. The design of our experimental study was built to accommodate these features.

The experiment consists of two parts, the main task (a forecasting task in the spirit of Grether, 1980) and a brief survey. In the forecasting task, subjects receive information about a risky asset, whose payoffs are either drawn from a "good distribution" or from a "bad distribution". Both distributions have three outcomes, which are identical across distributions but differ with respect to the probability with which they occur. All three outcomes can be ranked based on the payoff they generate and are thus labeled high, medium, or low. In the good distribution, the high payoff occurs with a 50 % probability, while the low payoff occurs with a 20% probability. In the bad distribution, probabilities are reversed, i.e. the low payoff occurs with a 50 % probability, while the high payoff occurs with only 20 % probability. Importantly, and crucial for the experimental design, the medium payoff always occurs with 30 % probability, irrespective of whether the distribution is good or bad. This ensures that the medium outcome does not provide any information about the underlying distribution, from which outcomes are drawn.

We introduce a 2x2 between-subject design with respect to the forecasting task. The first treatment dimension to which subjects are assigned depicts the potential payoffs of the two distributions. In particular, subjects are randomly assigned to either a *"positive"* condition or a *"negative"* condition. The three possible payoffs in the positive condition are +5 (high outcome), +1 (medium outcome), or -3 (low outcome). In the negative condition, all outcomes are shifted by -2 to keep the higher moments of the distribution constant while reducing the mean. As such, the three possible payoffs in the negative condition are +3 (high outcome), -1 (medium outcome), or -5 (low outcome). Table 1 provides a brief overview of possible outcomes across treatments.

[INSERT TABLE 1 ABOUT HERE]

The second treatment dimension relates to the set of questions subjects have to answer in the forecasting task. Subjects can be assigned to an "*active*" or a "*passive*" condition. In both conditions, subjects observe a payoff of the risky asset in ten consecutive rounds. Before the first round, the computer randomly determines the distribution of the risky asset (which can be good or bad). In the active condition, subjects decide before the beginning of each round whether they want to invest in the risky asset (whose distribution they have to forecast) or a bond, which always pays the medium outcome (-1, or +1; depending on the treatment) for sure. The payoff of the bond is thus equal to the expected value of the risky asset when no information about the underlying distribution is available. If the good distribution becomes more likely (i.e. occurs with a probability of greater than 50 %), the expected value of the risky asset is greater then the expected value of the bond, and vice versa. After their decision to invest in one of the two securities, subjects observe the payoff of the risky asset (irrespective of their choice) and are reminded of how much they have earned so far given their prior choices. Finally, subjects are asked to provide an estimate of the probability that the risky asset was paying from the good distribution and to rate their confidence in this estimate (assessed on a seven-point Likert scale). In the passive condition, subjects do not make any investment decision and start each round by observing the payoff of the risky asset in that trial. Afterwards, they are immediately asked to provide an estimate of the probability that the risky asset was paying from the good distribution and to rate their confidence in this estimate (also assessed on a seven-point Likert scale). An overview of all questions and the order in which subjects answer the questions is provided in Figure 1.

[INSERT FIGURE 1 ABOUT HERE]

To avoid potentially confounding factors resulting from biased memories (see Gödker et al., 2019), we explicitly display the prior outcomes in a table next to the questions. Additionally, we recognize that belief updating is an abstract task for many individuals. To ensure that subjects have a sufficient understanding of the forecasting task, they had to correctly answer three comprehension questions before they could continue (see Appendix A for the exact wording).

The experiment concluded with a brief survey about subjects' socio-economic background, self-assessed statistic skills, as well as a measure of risk preferences and financial literacy adopted from Kuhnen (2015). The latter two measures were obtained by asking subjects two questions regarding a portfolio allocation problem. In the first question, participants had to allocate \$10,000 between a broadly diversified index fund and a savings account. This answer provides a proxy for their risk preferences. The second question asked subjects to identify the correct formula for calculating the expected value of the portfolio they selected. Through multiple-choice answers, we can detect whether people lacked an understanding of probabilities, of the difference between net and gross returns, or of the difference between stocks and savings accounts. This yielded a financial knowledge score between zero to three (exact wording of questions is provided in the Appendix).

In the active condition participants were paid based on their investment payoffs and the accuracy of the probability estimates provided. Specifically, they received one twentieth of their accumulated payoffs (in the negative condition, all outcomes were shifted by +2 for the final calculation to make payment equivalent), plus 10 Cents for each probability estimate within 5 % of the objective Bayesian value. As such, subjects were motivated to be accurate in their forecasts and to maximize their payoffs. In the passive condition participants were paid based on the accuracy of the provided probability estimates, with the same rules as in the active condition.[5]

C. Hypotheses

To obtain parameter estimates for our main specification in the conceptual framework, we estimate a regression based on the natural logarithm of Eq. (1.3).

$$\ln\left(\frac{\lambda(G|s_1,\dots,s_t)}{\lambda(B|s_1,\dots,s_t)}\right) = \widehat{\beta_1}\cdot D_{informative} + \widehat{\beta_2}\cdot \ln\left(\frac{\lambda(G|s_1,\dots,s_{t-1})}{\lambda(B|s_1,\dots,s_{t-1})}\right) + \widehat{\beta_3} \cdot D_{uninformative\,|\,desirable} + \beta_4 \cdot D_{uninformative\,|\,undesirable} + \varepsilon_t. \tag{1.4}$$

Note that within a round *t*, the natural logarithm of subject's *i* odds ratio, based on her stated probability $P_{it}(G|s_1,\dots,s_t)$ that the asset is paying dividends from the good state is:

$$\lambda_{it} = \ln\left(\frac{\lambda(G|s_1,\dots,s_t)}{\lambda(B|s_1,\dots,s_t)}\right) = \ln\left(\frac{P_{it}(G|s_1,\dots,s_t)}{1-P_{it}(G|s_1,\dots,s_t)}\right)$$

[5] While the resulting payment for the passive condition was lower on average, participants also completed the experiment faster.

which may differ from the objective Bayesian probability. To make sure that the above ratio is defined for all observations, we truncated the data to lie in the [0.01, 0.99] interval. The interpretation of λ_{it} is straightforward. If λ_{it} is greater than (less than) zero, then person *i* beliefs in round *t* that the asset is more (less) likely to draw from the good state.

In the regression specification, we replaced $ln\frac{p(S|G)}{p(S|B)}$ with a dummy $D_{informative}$ taking the value 1 if the *t*th signal is g, 0 if the *t*th signal is u, and -1 if the *t*th signal is b (see Appendix C for a derivation). While this specification is equivalent (see Benjamin, 2019), we need to interpret the coefficient $\widehat{\beta_1}$ relative to $\left(\frac{\theta}{1-\theta}\right)$ instead of 1. Even though we have three possible outcomes in our experimental environment, which occur with 50 % (signal g or b), 30 % (signal u), and 20 % (signal g or b), only two of them are informative about the objective state of the world (signal g and signal b). Thus, the diagnosticity of an informative signal is set to $\theta = \frac{0.5}{0.5+0.2} = 0.714$ in our experiment and we need to interpret the coefficient $\widehat{\beta_1}$ relative to $\left(\frac{\theta}{1-\theta}\right) = \frac{0.714}{1-0.714} \approx 0.916.$

The regression specification allows us to control for several deviations from Bayesian behavior simultaneously, while testing whether individuals systematically incorporate uninformative signals into their beliefs. More precisely, if individuals are subject to conservatism (overreaction), one would expect $\widehat{\beta_1} < (>)\ 0.916$. Similarly, if individuals put too little (much) weight on their priors, one would expect $\widehat{\beta_2} < (>)\ 1$. Importantly, if people falsely incorporate uninformative signals in their belief formation process, one would observe that both $\widehat{\beta_3}$ and $\widehat{\beta_4}$ predict log-odds. In contrast, a test of Bayesian behavior would be:

$$\widehat{\beta_1} = \ln\left(\frac{\theta}{1-\theta}\right) = 0.916 \qquad \widehat{\beta_2} = 1 \qquad \widehat{\beta_3} = \widehat{\beta_4} = 0$$

D. Summary Statistics

Table 2 presents summary statistics. Overall, six-hundred forty-one individuals (420 males, 221 females, mean age 33 years, 8.8 years standard deviation) were recruited from Amazon Mechanical Turk (MTurk) to participate in an online experiment. MTurk advanced to a widely used and accepted recruiting platform for economic experiments. Not only does it offer a large and diverse subject pool compared to lab studies (which frequently rely on students), but it also provides a response quality similar to that of other subject pools (Buhrmester et al., 2011; Goodman et al., 2013). Participants reported average statistic skills of 4.71 out of 7 and would invest roughly 49 percent of their hypothetical endowment into a risky fund, which will serve as a proxy of risk aversion. Moreover, participants achieved a financial literacy score of approximately 1.34 out of 3.

[INSERT TABLE 2 ABOUT HERE]

Additionally, we tested whether the randomization from our between-subject design successfully resulted in a balanced sample. Table 2 also reports the mean and standard deviation for each control variable split by whether the uninformative signal was positive or negative (Panel A) and whether participants played the active or passive version of the forecasting task (Panel B). Differences were tested using rank-sum tests, or χ^2-tests for binary variables. Generally, we barely find any significant difference between our treatments, suggesting that our randomization was successful. The only exception are minor differences in financial literacy for the first treatment dimension (Panel A) and minor differences in self-reported statistic skills for the second treatment dimension (Panel B). We control for these factors in all our further analyses. For the remaining variables, we cannot reject the null hypothesis that the socio-economic background of the subjects is balanced between our treatments.

3 Results

3.1 Belief Updating after Uninformative Signals

Before we delve into the statistical analysis, Figure 2 visualizes participants general updating tendency after good, bad, and uninformative signals and compares it to Bayesian behavior. The figure displays results separately by whether the uninformative signal was positive (positive treatment) or negative (negative treatment).

[INSERT FIGURE 2 ABOUT HERE]

As can be inferred, subjects' beliefs adjust in the appropriate direction after both good and bad (informative) signals. Relative to Bayesian beliefs, we observe conservatism on average as subjects generally update too little both after good and after bad news. However, even after uninformative signals, subjects' beliefs adjust substantially and in the direction of the domain of the uninformative signal. While a Bayesian decision maker would not update his prior beliefs at all, subjects increase their priors after observing a positive uninformative signal, whereas they decrease their priors after observing a negative uninformative signal. Additionally, the strength with which they update their priors is symmetric for positive and negative uninformative signals.

While the pattern in Figure 2 provides first insights into subjects' updating behavior it is, of course, insufficient to justify a causal interpretation. To provide more formal evidence of how individuals update their priors after observing uninformative signals, we estimate OLS regressions of Eq. (1.4):

$$\lambda_{i,t} = \widehat{\beta_1} \cdot D_{informative;\, i,t} + \widehat{\beta_2} \cdot \lambda_{i,t-1} + \widehat{\beta_3} \cdot D_{uninformative;\, i,t} + \widehat{\beta_4} \cdot D_{informative;\, i,t}\ x\ negative_i + \sum_{j=1}^{n} \beta_j \boldsymbol{X}_{ij} + \varepsilon_{i,t}, \qquad (1.5)$$

where participants' subjective log-odds ratio is the dependent variable, and$\cdot$ $D_{informative;\, i,t}$ is a variable taking the value 1 if the *t*th signal is *good*, 0 if the *t*th signal is *uninformative*, and -1 if the *t*th signal is *bad*. $\lambda_{i,t-1}$ denotes the use of priors (i.e. the base-rate) and is defined as $\ln \frac{\lambda(G|s_1, \ldots, s_{t-1})}{\lambda(B|s_1, \ldots, s_{t-1})}$. $\cdot D_{uninformative;\, i,t}$ is a dummy if the *t*th signal of subjects i is uninformative, whereas $negative_i$ is a dummy if participant *i* is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant *i* encountered a negative uninformative signal in round *t*. Finally, $\boldsymbol{X}_{ij}$ is a set of control variables including age, gender, statistic skills, risk-aversion, and financial literacy. Results for the full sample and split by active and passive treatment are reported in Table 3.[6]

[INSERT TABLE 3 ABOUT HERE]

Our results suggest that Bayesian behavior is not predominant in the data. Specifically, we observe that $\widehat{\beta_1} \neq 0.915$, $\widehat{\beta_2} \neq 1$, and that both $\widehat{\beta_3}$ and $\widehat{\beta_4} \neq 0$. To interpret in which way

[6] Regression specifications are chosen to be identical to theory (i.e. estimated without a constant). However, other specifications yield similar results. We opt to present the simplest possible evidence.

individuals depart from Bayesian behavior, it is instructive to review what it would mean for individual coefficient estimates to vary from their Bayesian counterparts. Since $\widehat{\beta_1} < 0.915$ and $\widehat{\beta_2} < 1$, individuals suffer both from conservatism (i.e. they underinfer) and base-rate neglect (i.e. they under-use prior information). Most importantly however, we find that $\widehat{\beta_3} > 0$, whereas $\widehat{\beta_4} < 0$. In other words, even though a Bayesian would not update his prior beliefs after observing an uninformative signal, both positive and negative uninformative signals predict log-odds ratios. More precisely, controlling for both conservatism and base-rate neglect, individuals on average increase their priors after observing a positive uninformative signal with about half the strength as if the signal would contain information. Given the magnitude of both $\widehat{\beta_3}$ and $\widehat{\beta_4}$, this effect is mostly symmetric, suggesting that individuals also decrease their priors with about half the strength as if a negative uninformative signal would contain information. Lastly, we observe stronger effects when individuals have the opportunity to invest in the asset compared to when they simply state their beliefs. Taken together, this suggests that having stakes in the task exacerbates the bias resulting from uninformative signals, potentially because individuals *hope* to observe positive payoffs to maximize their earnings.

3.2 Frequency of updating mistakes

Thus far, we have established that individuals *on average* incorporate even uninformative signals into their beliefs based on the valence of the signal. However, these average patterns may mask a substantial amount of heterogeneity. In particular, it is not clear whether our results are driven by a few individuals who neglect the informational content but strongly focus on the valence or whether the here reported updating tendency applies to a large share of individuals, thus being a rather general phenomenon. To draw inference about the relation between the informational content and the valence of signals and to determine which aspect is most prevalent when processing uninformative signals, we examine how frequently individuals falsely update their beliefs after observing uninformative signals. To investigate the frequency, we define any belief update that is directionally inconsistent with the observed signal as an updating mistake. In the case of informative signals, an updating error is thus defined as a decrease (increase) in prior beliefs that the asset is drawing from the *good state* after subjects observed a good (bad) signal. Similarly, for uninformative signals, an updating error is defined as *any* update in prior beliefs after having observed an uninformative signal. Importantly, the definition above does not rely on the magnitude of the error, but only on the occurrence of such an error.

Figure 3 plots the absolute number of good, uninformative, and bad signals, as well as the number of mistakes after observing any of the three signals.

[INSERT FIGURE 3 ABOUT HERE]

Across all rounds and subjects, there are a total of 2,281 good signals, 1,929 uninformative signals, and 2,200 bad signals. Looking at informative signals, subjects only made basic errors in about 20% of the cases (18% and 23%, for good and bad signals, respectively). However, the rate at which subjects perform basic errors is substantially higher for uninformative signals. Here subjects updated their beliefs in 68% of the cases, even though the signal did not provide any learning opportunity about the underlying distribution. While Figure 3 already shows that the frequency of errors is substantially different for informative and uninformative signals, we further validate the robustness of the finding in a linear probability model. To do so, we estimate the following model[7]:

$$Error_{i,t} = \beta_0 + \beta_1 Uninformative_{i,t} + \beta_2 Objective\ Prior_{i,t} + \beta_3 Subjective\ Prior_{i,t} + Confidence_{i,t} + \sum_{j=1}^{n} \beta_j \boldsymbol{X}_{ij} + \varepsilon_{i,t},$$

where $Error_{i,t}$ is defined as individual *i* performing an updating error that is directionally inconsistent with the observed signal in round *t*. $Objective\ Prior_{i,t}$ is the rational prior for individual *i* as prescribed by Bayes' Theorem given the observed outcome history in round *t*, while $Subjective\ Prior_{i,t}$ is subjects' probability estimate in round *t*. Finally, $Confidence_{i,t}$ is subjects' self-reported confidence in their ability to provide correct probability forecasts. Results are reported in Table 4.

[INSERT TABLE 4 ABOUT HERE]

Consistent with our prior conjecture, we find that observing an uninformative signal in a given round substantially increases the likelihood of conducting an updating error. More precisely,

[7] While we estimate the model using OLS, results remain unchanged if we use a logit model instead.

we find that the likelihood of conducting an error is roughly 50 percentage points higher after observing an uninformative signal compared to observing a signal that does carry information about the underlying distribution. Interestingly, while this effect does not largely depend on the valence of the signal (Columns 3 and 4) it is less pronounced in the active treatment and more pronounced in the passive treatment (Column 2). The latter difference might be driven by the fact that subjects in the active treatment can derive payoff-relevant information from inferring the correct state of the underlying asset. As such, they might pay more attention to the information structure of the signals, thereby reducing their propensity to update their beliefs in response to uninformative signals. Besides the treatment, we find that the probability of updating one's beliefs in the wrong direction also correlates to participants' confidence in their own forecasts. Those individuals who are more confident that their forecast is correct are also less likely to update their beliefs in response to uninformative signals, suggesting that individuals are mindful about their own ability to provide correct forecasts.

Taken together, the analysis reinforces our prior evidence that individuals face difficulties in discerning the informational content of a signal from its valence. Additionally, the effect appears to be a general and quite robust phenomenon, as individuals more frequently update their beliefs after informationally irrelevant signals than they do not.

3.3 Mechanism: Reference-dependent belief updating

One common and persistent finding so far is that individuals not only face difficulties in correctly identifying the informational content of signals but also that they struggle to discern it from the valence of the signal. In this section, we explore potential mechanisms underlying this pattern.

To do so, we focus on the influence of participants' prior beliefs about the state of the risky asset. Prior beliefs have previously been shown to affect updating mistakes and biased inference in multiple ways and thus serve as a natural starting point for our analysis. Testing the implications of a model by Rabin and Schrag (1999), both Charness and Dave (2017) and Pouget et al. (2017) find evidence that individuals draw inference in a manner that is biased in favor of current beliefs about the objective state of the risky asset. A related body of research documents that people update their beliefs about future outcomes in an asymmetric manner: they tend to neglect undesirable information, and overweight desirable information (Eil & Rao, 2011; Möbius et al., 2014; Sharot & Garret, 2016). In our experiment, prior beliefs are important for two reasons. First, when deciding between investing in the risky asset or choosing the risk-free alternative, holding a particular belief has direct consequences for the investment decision. As such, beliefs have a value in and of themselves, as positive beliefs about the state of the risky asset are related to higher potential payoffs. Second, and perhaps even more important, extreme priors (both optimistic and pessimistic) are usually the result of observing one particular signal more frequently than the other signals (i.e. very optimistic beliefs usually develop in response to observing many *good* signals). In our environment, good signals are always associated with the highest payoff, whereas bad signals are associated with the lowest payoff and uninformative signals with a medium payoff (as illustrated in Table 1). Thus, extreme priors (which develop in tandem with the associated high, medium, or low payoffs) might shift participants' reference point. To illustrate this idea, consider a participant who frequently observes the *good* signal with the respective high payoff. Such a participant might react differently when observing a positive uninformative signal (with medium payoffs) compared to someone who frequently observes *bad* signals.

To differentiate optimistic from pessimistic priors, we define a prior that the asset is drawing from the good distribution greater than 50 percent as positive prior and a prior that the asset is drawing from the good distribution smaller than 50 percent as negative prior.[8] Figure 4 visualizes participants updating behavior after non-diagnostic signals, split by treatment (positive vs. negative) and by prior.

[INSERT FIGURE 4 ABOUT HERE]

Figure 4 reveals that priors appear to play an important role in processing uninformative signals. Those subjects who hold positive priors only update their beliefs weakly in response to positive uninformative signals, whereas they update strongly in response to negative uninformative signals. Symmetrically, subjects who hold negative priors substantially increase their priors after observing positive uninformative signals, whereas they only weakly update their priors after observing negative uninformative signals. This pattern suggests that not only the domain

[8] This definition is consistent with the point where one of the two assets has a higher expected value. For priors greater (smaller) than 50 percent, the expected value of the risky asset is greater (smaller) than the expected value of the riskless security.

of the uninformative signal is important but also how desirable the signal is in relation to what subjects expect to observe.

To more rigorously investigate how prior beliefs affect our previously reported results, we estimate OLS regressions of our main specification (Eq. 1.4) split by prior beliefs and by whether participants are invested in the risky asset or not. Importantly, we only include participants from the active treatment in the analysis, as the decision to be invested in the asset or not is most likely a deliberate choice that depends on prior beliefs.[9] Table 5 reports results.

[INSERT TABLE 5 ABOUT HERE]

Coefficient estimates for actively invested participants reveal a fundamental asymmetry in how the processing of uninformative signal depends on prior beliefs. Those subjects who hold optimistic prior beliefs about the state of the risky asset (i.e. subjects who belief the good outcome is more likely to occur) only weakly increase their beliefs when the uninformative signal is positive (but in magnitude smaller than the good signal), but strongly decrease their beliefs when the signal is negative. Similarly, subjects who hold pessimistic prior beliefs about the state of the risky asset only weakly increase their beliefs when the uninformative signal is negative (but in magnitude smaller than the bad signal), but strongly increase their beliefs when the signal is positive. As such, subjects appear to process uninformative signals not exclusively on the basis of the valence of the signal, but rather relative to some reference point which is dictated by their prior beliefs.

Importantly, this finding cannot be explained by prior-biased inference (or confirmation bias) as tested by Charness and Dave (2017) and Pouget et al. (2017) as individuals show stronger reactions to the valence uninformative signals that contradict their prior beliefs. Additionally, this finding is also different from preference-biased inference (Eil & Rao, 2011; Möbius et al., 2014) as individuals also overreact to undesirable signals. Subjects both strongly react to positive uninformative signals that contradict pessimistic priors as well as to negative uninformative signals that contradict optimistic priors. Instead, it appears that subjects incorporate uninformative signals in a reference-dependent manner, dictated by their prior beliefs. They fail to correctly identify that uninformative signals do not carry information about

[9] Results for participants in the passive treatment are directionally consistent. However, we decide to present the results that are undoubtedly affected by participants' prior beliefs.

the objective state of the world and update their beliefs based on the valence of the signal relative to their current prior expectations. Yet, given that we find that individuals react strongest to desirable uninformative signals when their priors are pessimistic, it appears that they seek to revert very pessimistic priors most quickly, consistent with the model of Bénabou (2013).

However, similar conclusions cannot be drawn for subjects who are not actively invested in the risky asset. While those who hold positive priors update their beliefs rather symmetrically in the direction of the domain of the uninformative signal, individuals who hold negative priors do not appear to update their beliefs at all after uninformative signals. One potential explanation for this behavior is that subjects who are rather optimistic about the state of the risky asset (while not being invested) still follow the outcomes to invest in a future round once they become more certain of the state. Finally, individuals who are not invested and hold pessimistic beliefs might simply not pay enough attention to the outcomes, as they continue to collect the risk-free payoff.

Taken together, our results suggest that being invested in the risky asset appears to be a necessary condition for subjects to engage in reference-dependent updating following uninformative signals. More generally, there has to be some intrinsic or extrinsic advantage for holding a particular belief such as making a payoff-relevant investment decision.

3.4 Robustness Checks

A. The role of memory and learning

Two important concepts related to the formation of probabilistic beliefs are the role of memory and learning effects. However, our experiment was constructed in a way to ensure that neither of the effects can account for our findings. First, subjects are always provided with the full outcome history of prior signals. In particular, as can be seen in the Appendix, the history of prior signals is clearly displayed next to the forecasting question. Additionally, our experimental design does not provide feedback and hence little scope for learning. Moreover, it is highly doubtful that subjects would learn within the course of ten experimental periods even in the presence of feedback. To verify that the effect is not driven by initial forecasting errors when subjects lack the experience and potentially less pronounced the more signals individuals observe, we separately estimate our main specification for the first five and the final five signals that individuals observe. Table B1 in the Appendix reveals that coefficient estimates for positive and negative uninformative signals remain relatively stable throughout the experiment. Together with the fact that we did not provide any feedback, we conclude that the effect appears to be stable over time.

B. Sample splits

Finally, we replicate our main analyses on different subsamples to validate its robustness. In particular, we conduct splits regarding (i) extreme outliers; (ii) "speeders"; and (iii) forecasting performance. Extreme outliers are individuals whose subjective priors largely deviate from the Bayesian benchmark and who frequently update in the wrong direction. Similar to the exclusion criteria of Enke and Graeber (2019), we define extreme outliers as individuals who report a subjective posterior $p_s < 25\ \% \ (> 75\ \%)$ when the Bayesian posterior is $p_B > 75\ \% \ (< 25\ \%)$. Speeders are defined as subjects who are in the bottom quintile of the response time distribution. Finally, we also conduct splits regarding how subjects overall performed in the forecasting task. To verify that the effect does not capture those individuals who showed difficulties in understanding the task, we define the squared deviation of subjects' probability estimate in each period from the objective posterior probability as a measure of forecasting quality and conduct median splits. Results are reported in Table 6.

[INSERT TABLE 6 ABOUT HERE]

Overall, results are very similar across all subsamples and confirm our previously drawn conclusions. First, we consistently find that uninformative signals predict log-odds ratios in every subsample. Second, similar to our main analysis, positive uninformative signals predict an increase in the log-odds ratio, whereas negative uninformative signals predict a decrease, with the effect being of similar strength. Lastly, we also find differences between the different subgroups. In particular, extreme outliers, speeders and individuals with below-median forecasting performance show more pronounced effects both for positive and negative uninformative signals.

4 Conclusion

This article experimentally studies how individuals update their beliefs after observing non-diagnostic information signals with varying degrees of desirability. Whereas Bayes' Rule predicts that such *uninformative* signals do not influence inference judgements, we find that individuals systematically incorporate them in their belief formation process. Importantly, the direction in which individuals update their beliefs strongly depends on the valence of the observed signal. Individuals tend to form more optimistic beliefs about the objective state of the world after observing desirable uninformative signals, whereas they form more pessimistic beliefs after observing undesirable uninformative signals. As mechanism we identify that individuals process the valence of new signals in a reference-dependent manner, dictated by their prior beliefs. Whenever they observe non-diagnostic outcomes which are close to their prior expectations, they only weakly update their beliefs, whereas when they observe non-diagnostic outcomes which are at odds with their prior expectations, they strongly overreact.

Taken together, our findings suggest that individuals appear to struggle discerning belief-relevant information from their preferences. Such deviations from Bayesian behavior are particularly severe in situations in which the valence of non-diagnostic signals is at odds with the valence of objective pieces of information. Even though decision making frequently involves the accumulation of new pieces of information until uncertainty is reduced to a tolerable level, such a bias may instead lead to a decline in predictive performance.

References

Bénabou, R. (2013). Groupthink: Collective delusions in organizations and markets. *Review of Economic Studies*, *80*(2), 429-462.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 2, pp. 69-186). North-Holland.

Benoît, J. P., & Dubra, J. (2018). When do populations polarize? An explanation. Working Paper.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science*, *6*(1), 3-5.

Charness, G., & Dave, C. (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior*, *104*, 1-23.

Dave, C. and Wolfe, K. W. (2003). On confirmation bias and deviations from bayesian updating. Working Paper.

Eddy, D.M., 1982. Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman, D., Slovic, P., Tversky, A. (Eds.), Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press, New York, pp. 249–267.

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114-38.

Enke, B., & Graeber, T. (2019). *Cognitive Uncertainty* (No. w26518). National Bureau of Economic Research. Working Paper.

Fischhoff, B., & Bar-Hillel, M. (1984). Diagnosticity and the base-rate effect. *Memory & Cognition*, 12(4), 402-410.

Ginosar, Z., Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of Experimental Social Psychology* 16, 228–242.

Ginosar, Z., Trope, Y. (1987). Problem solving in judgment under uncertainty. Journal of *Personality and Social Psychology* 52 (3), 464–474.

Gödker, K., Jiao, P., & Smeets, P. (2019). Investor Memory. Working Paper.

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, 26(3), 213-224.

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

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection: An experimental study. *Judgment and Decision making*, 8, 407-24.

Kuhnen, C. M. (2015). Asymmetric learning from financial information. *The Journal of Finance*, 70(5), 2029-2062.

Lord, C.G., Ross, L., & Lepper, M.R. (1979). Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37 (11), 2098–2109.

Meehl, P.E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin* 52 (3), 194–216.

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2014). Managing self-confidence. NBER Working Paper.

Phillips, L.D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology* 72 (3), 346–354.

Pouget, S., Sauvagnat, J., & Villeneuve, S. (2017). A mind is a terrible thing to change: confirmatory bias in financial markets. *The Review of Financial Studies*, 30 (6), 2066-2109.

Rabin, M., & Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114 (1), 37-82.

Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in cognitive sciences*, 20(1), 25-33.

Sharot, T., Korn, C.W., & Dolan, R.J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience* 14 (11), 1475–1479.

Tribe, L.H. (1971). Trial by mathematics: precision and ritual in the legal process. *Harvard Law Review* 84 (6), 1329–1393.

Troutman, C.M., & Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organizational Behavior and Human Performance* 19 (1), 43–55.

Tversky, A., & Kahneman, D. (1971). Belief in the Law of Small Numbers. *Psychological Bulletin* 76 (2), 105–110.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185 (4157), 1124–1131.

Wells, G.L., & Harvey, J.H. (1978). Naïve attributors' attributions and predictions: what is informative and when is an effect an effect? *Journal of Personality and Social Psychology* 36 (5), 483–490.

Zimmerman, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110 (2), 337 – 361.

Figures

Figure 1: Overview of the Updating Task

This figure provides an overview of the questions subjects have to answer in the forecasting task. Subjects in the passive treatment have to answer three questions in each round, whereas subjects in the active treatment have to answer five questions in each round (denoted with [active only]). Subjects make forecasting decisions in 10 consecutive rounds.

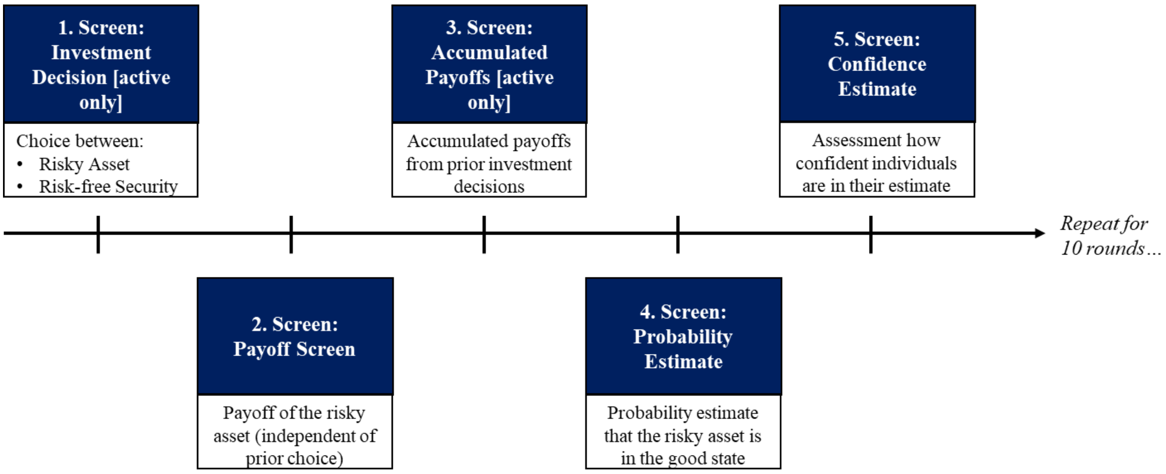


Figure 2: General Updating Tendency

This figure illustrates subjects' general belief updating after observing good, uninformative, and bad signals about the state of the risky asset. Displayed are actual changes in prior beliefs as well as the correct Bayesian change in probability. Results are displayed separately by whether subjects encountered the uninformative signal in the positive or negative domain. Displayed are 95% confidence intervals.

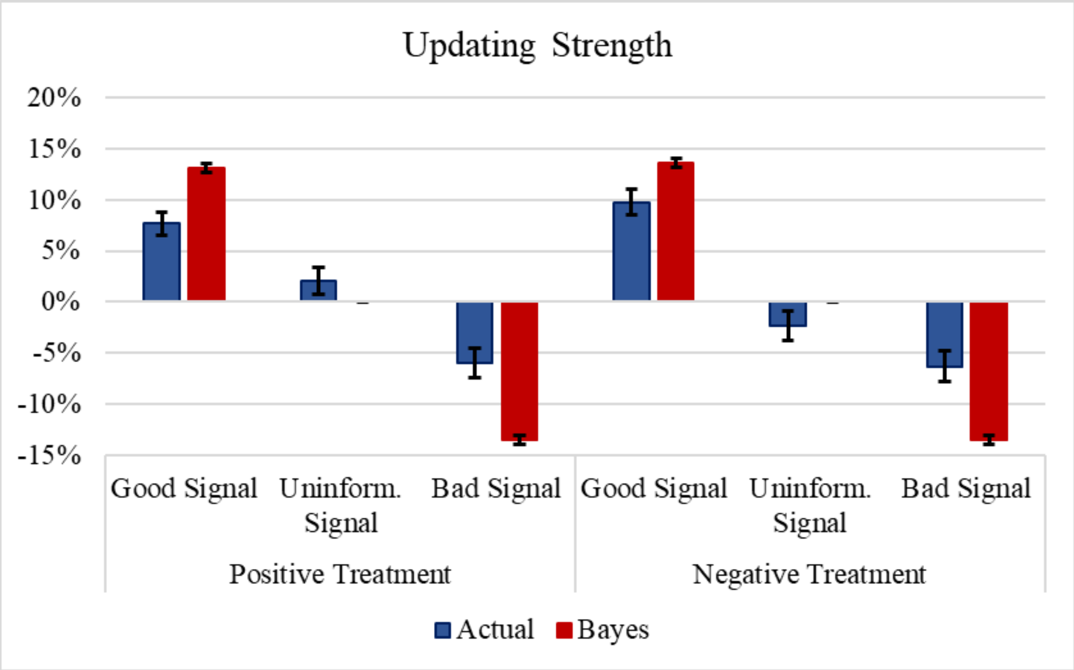


Figure 3: Basic Updating Mistakes

This figure illustrates the number of directionally inconsistent belief updates relative to the overall number of observed signals. Results are displayed separately for good, uninformative, and bad signals.

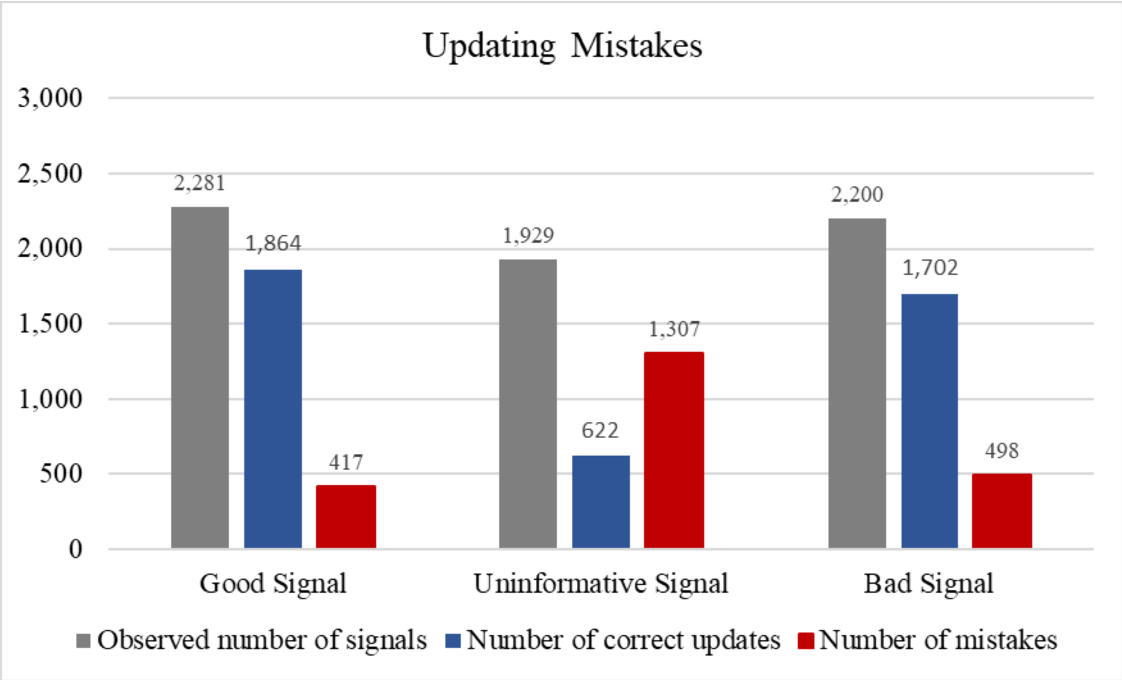
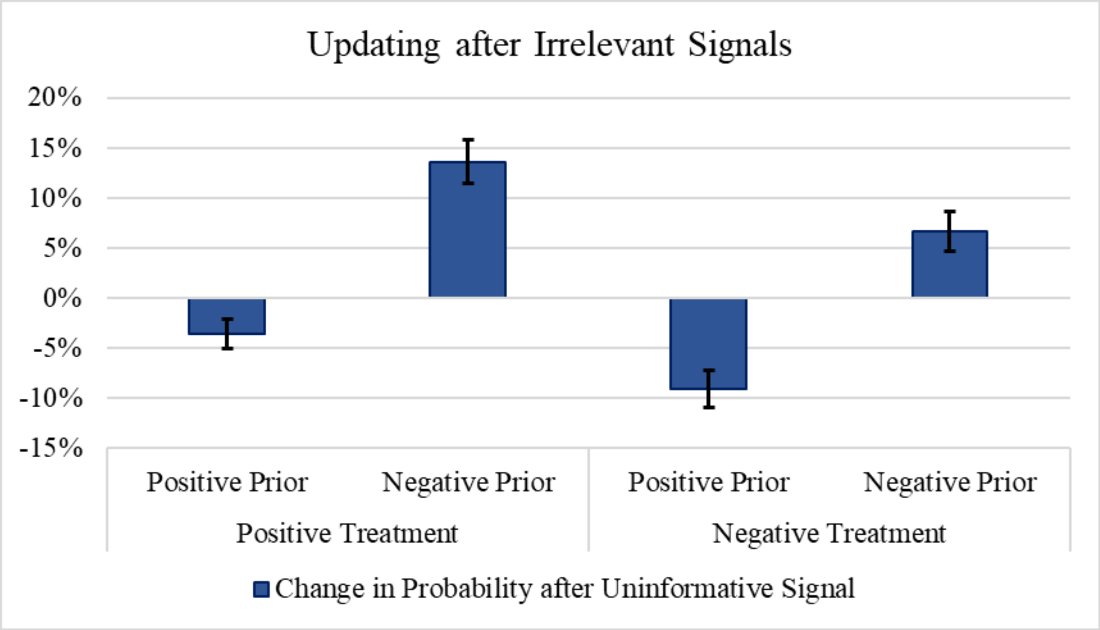


Figure 4: General Updating Tendency

This figure illustrates the change in prior beliefs after observing uninformative signals split by positive and negative prior beliefs and by treatment. Results are displayed separately for the positive and negative treatment. Displayed are 95% confidence intervals.



Tables

Table 1: Payoff Distribution

This table reports the payoffs associated with good, uninformative, and bad signals, split by positive and negative treatment.

Payoffs	**Good Signal**	**Uninformative Signal**	**Bad Signal**
Positive Treatment	+5	+1	-3
Negative Treatment	+3	-1	-5

Table 2: Summary statistics

This table shows summary statistics for our experimental data. Reported are the mean and the standard deviation (in parentheses) for the whole sample (Column 1) and split across treatments (Panel A for Positive/Negative; Panel B for Active/Passive). Column 4 presents randomization checks. Differences in mean were tested using rank-sum tests, or χ^2-tests for binary variables. The p-value is reported in Column 5. *Female* is an indicator variable that equals 1 if a participant is female. *Statistic skills* denotes participants' self-assed statistical skills on a 7-point Likert scale. *Risk Preferences* is the percentage of their initial endowment that subjects invested in a risky investment option. *Financial Literacy* is a score between zero (lacking basic understanding) to three (correctly answered each question) as defined in Section 2.

Panel A **Variable**	**Full sample (N = 641)**	**Positive Treatment (N = 321)**	**Negative Treatment (N = 320)**	**Difference**	**p-value**
Age	33.00	33.00	33.00	0.00	1.00
	(8.79)	(9.04)	(8.55)		
Female (1 = Yes)	0.34	0.36	0.33	0.03	0.38
	(0.48)	(0.48)	(0.47)		
Statistic Skills (1-7)	4.71	4.76	4.66	0.10	0.47
	(1.76)	(1.73)	(1.78)		
Risk Preferences (% invested in	0.49	0.49	0.50	0.01	0.51
risky asset)	(0.31)	(0.30)	(0.31)		
Financial Literacy	1.34	1.43	1.26	0.17	0.01
	(0.91)	(0.88)	(0.93)		

Panel B **Variable**	**Full sample (N = 641)**	**Passive Treatment (N = 330)**	**Active Treatment (N = 311)**	**Difference**	**p-value**
Age	33.00	33.49	32.48	1.01	0.15
	(8.79)	(8.47)	(9.11)		
Female (1 = Yes)	0.34	0.36	0.32	0.04	0.30
	(0.48)	(0.48)	(0.47)		
Statistic Skills (1-7)	4.71	4.55	4.87	0.31	0.02
	(1.76)	(1.83)	(1.67)		
Risk Preferences (% invested in	0.49	0.48	0.51	0.03	0.18
risky asset)	(0.31)	(0.28)	(0.33)		
Financial Literacy	1.34	1.35	1.34	0.01	0.85
	(0.91)	(0.92)	(0.89)		

Table 3: Main Result

This table reports the results of three OLS regressions on how information signals and their valence affect individuals' beliefs. We report results for the full sample and split by active and passive treatment. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $D_{informative;i,t}$ is a variable taking the value 1 if the *t*th signal of subject *i* is *good*, 0 if the *t*th signal is *uninformative*, and -1 if the *t*th signal is *bad*. $D_{uninformative;i,t}$ is a dummy if the *t*th signal of subject *i* is uninformative, whereas $negative_i$ is a dummy if participant *i* is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant i encountered a negative uninformative signal in round t. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	*Log Odds Ratio (Subjective)* $\lambda_{i,t}$		
	Full Sample	Active	Passive
$D_{informative;i,t}$ *(Inference)*	0.472***	0.369***	0.569***
	(15.22)	(9.06)	(12.39)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.697***	0.758***	0.624***
	(32.99)	(27.91)	(19.60)
$D_{uninformative;i,t}$	0.256***	0.329***	0.152**
	(5.89)	(5.16)	(2.56)
$D_{uninformative;i,t}$ x $negative_i$	-0.514***	-0.563***	-0.443***
	(-7.19)	(-5.40)	(-4.54)
Observations	5769	2799	2970
R^2	0.538	0.611	0.468

Table 4: Frequency of directionally inconsistent updating errors

This table reports the results of four OLS regressions on how frequently individuals perform directionally inconsistent updating mistakes. We report results for the full sample and split by positive and negative treatment. The dependent variable is *Updating Error*, a dummy that equals 1 if participants perform a updating mistake that is directionally inconsistent with Bayes' Rule as defined in Section 3.2. $D_{uninformative;i,t}$ is a variable taking the value 1 if the *t*th signal is *uninformative*, and 0 otherwise. *Objective Posterior* is the correct Bayesian probability that the risky asset is in the good state, given the information seen by the participant up to trial *t* in the learning block. *Subjective Probability Estimate* and *Confidence Estimate* are participants' estimates of the probability that the risky asset is in the good state and their assessed confidence, respectively. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	*Updating Error*			
	Full Sample	Full Sample	Positive Treatment	Negative Treatment
$D_{uninformative;i,t}$	0.475***	0.517***	0.449***	0.506***
	(33.42)	(26.54)	(22.35)	(25.08)
Objective Posterior	-0.00046	-0.00046*	-0.00012	-0.00079**
	(-1.63)	(-1.67)	(-0.30)	(-2.10)
Subjective Probability Estimate	0.00067**	0.00077**	-0.00010	0.0014***
	(1.99)	(2.31)	(-0.20)	(3.17)
Confidence Estimate	-0.0147***	-0.0153***	-0.0111*	-0.0166***
	(-3.71)	(-3.84)	(-1.81)	(-3.15)
Active		-0.00983		
		(-0.59)		
$D_{uninformative;i,t}$ x *Active*		-0.0850***		
		(-3.03)		
Constant	0.255***	0.256***	0.275***	0.241***
	(5.34)	(5.37)	(3.90)	(3.63)
Observations	6410	6410	3210	3200
R^2	0.224	0.227	0.204	0.250

Table 5: Mechanism: Reference-dependent belief updating

This table reports the results of four OLS regressions on how information signals and their valence affect individuals' beliefs. We report results split by whether participants are actively invested in the risky asset or not and by participants' prior beliefs about the state of the risky asset as defined in Section 3.3. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $D_{informative;i,t}$ is a variable taking the value 1 if the *t*th signal of subject *i* is *good*, 0 if the *t*th signal is *uninformative*, and -1 if the *t*th signal is *bad*. $D_{uninformative;i,t}$ is a dummy if the *t*th signal of subject *i* is uninformative, whereas $negative_i$ is a dummy if participant *i* is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant *i* encountered a negative uninformative signal in round *t*. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	*Log Odds Ratio (Subjective)* $\lambda_{i,t}$			
	Actively Invested		Not Invested	
	Positive Prior	Negative Prior	Positive Prior	Negative Prior
$D_{informative;i,t}$ *(Inference)*	0.354*** (7.13)	0.459*** (6.01)	0.214*** (2.81)	0.592*** (4.85)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.797*** (22.55)	0.755*** (11.04)	0.748*** (14.88)	0.814*** (9.77)
$D_{uninformative;i,t}$	0.213** (2.38)	0.740*** (3.52)	0.307** (2.51)	0.221 (1.00)
$D_{uninformative;i,t}$ x $negative_i$	-0.579*** (-3.58)	-0.842*** (-3.21)	-0.545** (-2.52)	-0.302 (-0.79)
Observations	1371	530	533	287
R^2	0.698	0.454	0.580	0.538

Table 6: Robustness Checks

This table reports the results of six OLS regressions to investigate the robustness of our main finding. We report sample splits based on three measures as defined in Section 3.4: (1) strong outliers; (2) speeder; and (3) forecast quality. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $D_{informative;i,t}$ is a variable taking the value 1 if the *t*th signal of subject *i* is *good*, 0 if the *t*th signal is *uninformative*, and -1 if the *t*th signal is *bad*. $D_{uninformative;i,t}$ is a dummy if the *t*th signal of subject *i* is uninformative, whereas $negative_i$ is a dummy if participant *i* is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant *i* encountered a negative uninformative signal in round *t*. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	*Log Odds Ratio (Subjective)* $\lambda_{i,t}$					
	(1) Outlier?		(2) Speeder?		(3) Above Median Forecaster?	
	No	Yes	No	Yes	No	Yes
$D_{informative;i,t}$ *(Inference)*	0.561*** (17.54)	0.168** (2.41)	0.493*** (15.69)	0.259** (2.02)	0.263*** (5.48)	0.610*** (16.71)
$\lambda_{i,t-1}$ *(Use of Priors)*	0.734*** (35.29)	0.533*** (10.32)	0.715*** (33.48)	0.547*** (6.83)	0.545*** (14.74)	0.800*** (41.65)
$D_{uninformative;i,t}$	0.238*** (4.88)	0.383*** (4.70)	0.226*** (5.23)	0.702*** (3.41)	0.418*** (5.99)	0.161*** (3.48)
$D_{uninformative;i,t}$ x $negative_i$	-0.444*** (-6.09)	-0.701*** (-4.18)	-0.436*** (-6.15)	-1.296*** (-4.31)	-0.716*** (-6.43)	-0.343*** (-4.47)
Observations	4275	1494	5193	576	2880	2889
R^2	0.638	0.340	0.557	0.477	0.747	0.344

Appendix

A. Experimental Instructions and Screenshots

Instructions Bayesian Updating (exemplary for positive treatment)

In this part we would like to test your forecasting abilities.

You will make forecasting decisions in ten consecutive rounds.

Suppose you find yourself in an environment, in which a risky asset can pay a dividend of either -3, +1, or +5. The probability of each outcome depends on the state in which the asset is (**good** state or **bad** state). If the risky asset is in the **good** state, then the probability that it pays a dividend of +5 is 50%, the probability that it pays a dividend of -1 is 30% and the probability that it pays a dividend of -3 is 20%. If the risky asset is in the **bad** state, then the probability that it pays a dividend of +5 is 20%, the probability that it pays a dividend of -1 is 30% and the probability that it pays a dividend of -3 is 50%.

The computer determines the state of the risky asset before the first round. Afterwards, the state does not change and remains fixed. At first, you do not know which state the risky asset is in. The risky asset may be in the good state or in the bad state with equal probability.

At the beginning of each round, you will observe a dividend payment of the risky asset (-3, +1, or +5). After that, we will ask you to provide a probability estimate that the risky asset is in the good state and ask you how sure you are about your probability estimate. While answering these questions, you can observe the previous dividend payments next to the question.

There is always an objective correct probability that the risky asset is in the good state. This probability depends on the history of payoffs of the risky asset already. As you observe the payoffs of the risky asset, you will update your beliefs whether or not the risky asset is in the good state.

Objective Bayesian Posterior Probabilities

This table provides all possible values for the objectively correct probability that the asset is in the good state for every possible combination of trials and outcomes. The initial prior for good and bad distribution is set to 50%. The objective Bayesian posterior probability that the asset is in the good state, after observing t high outcomes in n trials so far is given by: $\frac{1}{1+\frac{1-p}{p}\cdot(\frac{q}{1-q})^{n-2t}}$, where p is the initial prior before any outcome is observed that the stock is in the good state (50% here), and q is the probability that the value increase of the asset is the higher one.

Screenshots of the Experiment

Figures A1 to A5 present the screens of the forecasting task (screen that only belong to the active treatment are marked as [active]) as seen by subjects in the experiment. One round consists of three [five] sequential screens.

Figure A1: Investment screen [active only]

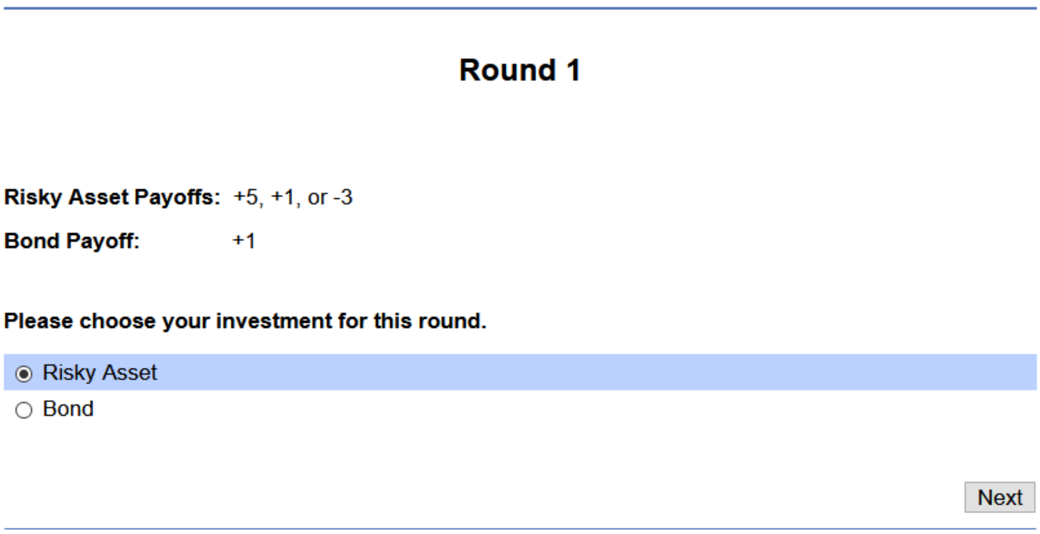


Figure A2: Payoff screen

Figure A3: Accumulated payoffs screen [active only]

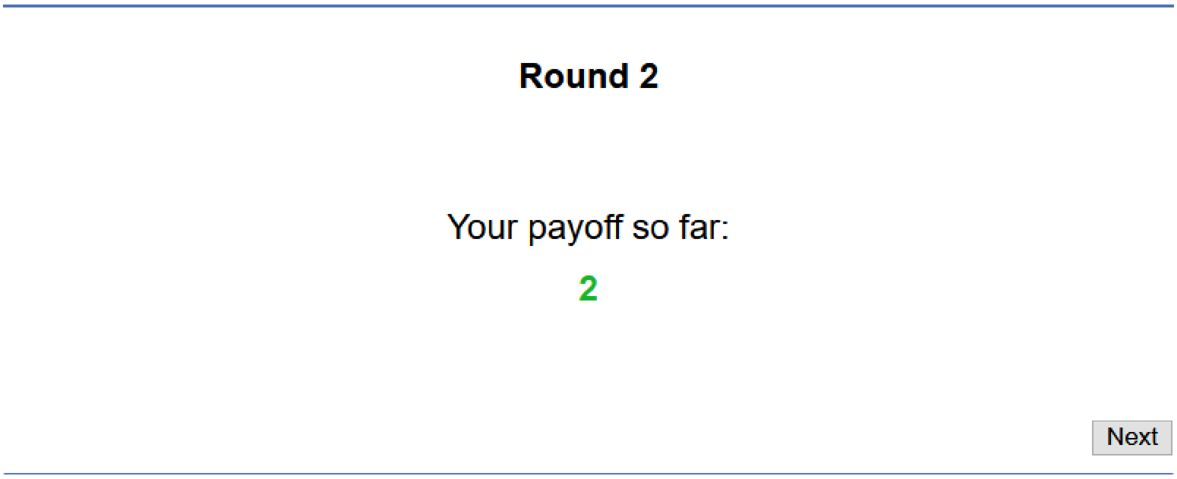


Figure A4: Probability estimate screen

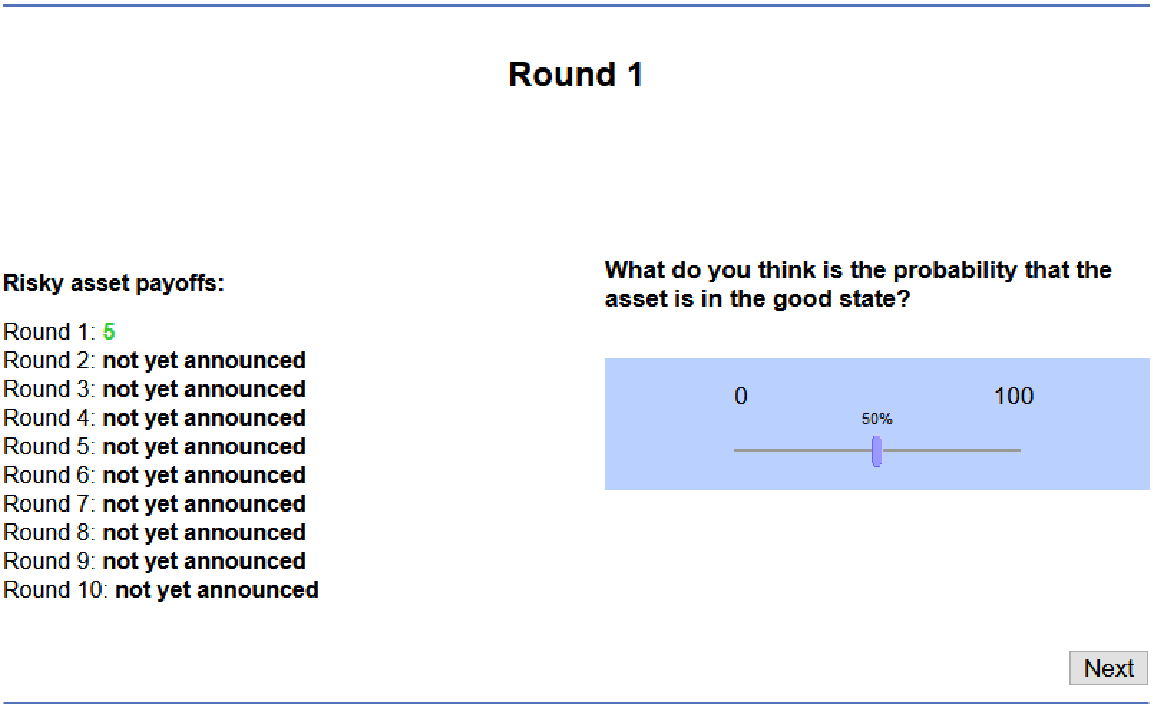
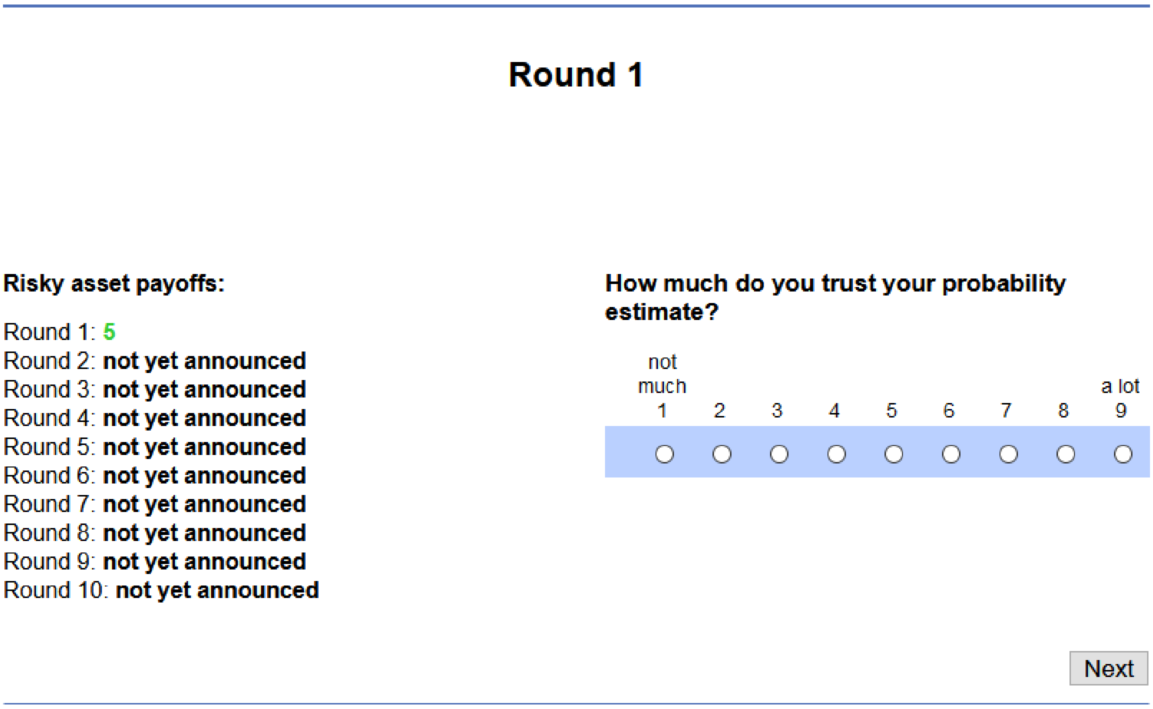


Figure A5: Confidence level screen



Comprehension Question for Bayesian Updating Task

Below we report the comprehension questions that participants had to answer correctly after reading the instructions to proceed to the Bayesian Updating task. Correct responses are displayed in *italic*.

1. If you see a series of +5, what is more likely?
    a. *The risky asset is in the good state.*
    b. The risky asset is in the bad state.

2. You observe a -5, how do you have to update your probability estimate that the asset draws from the good distribution?
    a. *I reduce the probability estimate that the asset is in the good distribution.*
    b. I increase the probability estimate that the asset is in the good distribution.

3. The correct probability estimate is let's say 0.70. Which probability estimate(s) would be in the range such that you earn 25 cents? [Note: You can check multiple boxes.]
    a. 0.55
    b. *0.67*
    c. *0.75*
    d. 0.85
    e. 0.87

4. At the beginning of the first period, the probability that the risky asset is in the good state is 50%.
    a. *True*
    b. False

Risk Aversion Question

Below we report the risk aversion question adopted from Kuhnen (2015):

Imagine you have saved $10,000. You can now invest this money over the next year using two investment options: a U.S. stock index mutual fund, which tracks the performance of the U.S. stock market, and a savings account. The annual return per dollar invested in the stock index fund will be either +40% or -20%, with equal probability. In other words, it is equally likely that for each dollar you invest in the stock market, at the end of the one year investment period, you will have either gained 40 cents, or lost 20 cents. For the savings account, the known and certain rate of return for a one year investment is 5%. In other words, for each dollar you put in the savings account today, for sure you will gain 5 cents at the end of the one year investment period. We assume that whatever amount you do not invest in stocks will be invested in the savings account and will earn the risk-free rate of return.

Given this information, how much of the $10,000 will you invest in the U.S. stock index fund? Choose an answer that you would be comfortable with if this was a real-life investment decision.

[Please enter a value between 0 and 10,000 here]

Financial Literacy Question

Below we report the financial literacy question adopted from Kuhnen (2015). Correct responses are displayed in *italic*.

Let's say that when you answered the prior question you decided to invest x dollars out of the $10,000 amount in the U.S. stock index fund, and therefore you put (10,000 - x) dollars in the savings account. Recall that over the next year the rate of return of the stock index fund will be +40% or -20%, with equal probability. For the savings account, the rate of return is 5% for sure. What is the amount of money you expect to have at the end of this one year investment period?

Please choose one of the answers below

[A] 0.5 (0.4x - 0.2x) + 0.05 (10,000 - x)

[B] 1.4x + 0.8x + 1.05 (10,000 - x)

[C] 0.4 (10,000 - x) - 0.2 (10,000 - x) + 0.05x

[D] 0.5 (0.4 (10,000 - x) - 0.2 (10,000 - x)) + 0.05x

[E] 0.4x - 0.2x + 0.05 (10,000 - x)

[F] 0.5 (1.4x + 0.8x) + 1.05 (10,000 - x)

[G] 1.4 (10,000 - x) + 0.8 (10,000 - x) + 1.05x

[H] 0.5 (1.4 (10,000 - x) + 0.8 (10,000 - x)) + 1.05x

B. Further Analyses

Table B1

This table reports the results of four OLS regressions on how information signals and their valence affect individuals' beliefs. To investigate the effect of learning, we report results split by the first five and the last five rounds of the experiment. The dependent variable is participants' subjective log-odds ratio as defined in Section 2. $Prior\ Signal_{i,t}$ is a variable taking the value 1 if the *t*th signal of subject *i* is *good*, 0 if the *t*th signal is *uninformative*, and -1 if the *t*th signal is *bad*. $Uninformative_{i,t}$ is a dummy if the *t*th signal of subject *i* is uninformative, whereas $negative_i$ is a dummy if participant *i* is in the negative treatment (and zero otherwise). The interaction term thus displays whether participant *i* encountered a negative uninformative signal in round *t*. Controls include age, gender, statistical skills, risk aversion, and participants' financial literacy. Reported are coefficients and t-statistics (in parentheses). All standard errors are clustered at the individual level. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	*Log Odds Ratio (Subjective)*			
	Active Treatment		Passive Treatment	
	Round 1 to 5	Round 6 to 10	Round 1 to 5	Round 6 to 10
Prior Signal	0.412***	0.340***	0.604***	0.536***
	(8.39)	(6.93)	(10.58)	(10.36)
Prior Beliefs	0.701***	0.789***	0.520***	0.689***
	(19.51)	(26.23)	(11.92)	(21.88)
Uninformative	0.367***	0.310***	0.257***	0.0831
	(4.16)	(3.93)	(2.98)	(1.19)
Negative	0.181***	0.182***	0.135**	0.0920
	(2.73)	(3.06)	(2.20)	(1.64)
Uninformative *x* Negative	-0.544***	-0.589***	-0.596***	-0.348***
	(-4.04)	(-4.11)	(-4.24)	(-2.88)
Observations	1371	530	533	287
R^2	0.698	0.454	0.580	0.538

C. Derivations and Proofs

Sequential Bayesian Updating Behavior

Following Dave and Wolfe (2003) and Charness and Dave (2017), we briefly sketch individuals' sequential updating behavior as prescribed by Bayes' Law:

Suppose there are two possible states of the world, denoted 'G' (for good) and 'B' (for bad). Additionally, over the course of *t* rounds, individuals may observe signals that are either indicative of a good state g, or of a bad state *b*, or which are non-diagnostic about the underlying state u. Within a given round t, Bayes' Rule assumes that individuals posterior logs π_{1k} are formed as a function of their prior logs π_0 and some likelihood L_k:

$$\pi_{1k} = L_k \pi_0$$

Given that only two signals are indictive about the possible states 'G' and 'B', the likelihood L_k takes the following form:

$$L_k = \left(\frac{\theta}{1-\theta}\right)^{z_t},$$

where θ is the proportion of g signals to *b* signals and z_t is the difference between the number of *g* signals and *b* signals as of the *t*th round. Note that only the difference of informative signals is important for the likelihood function, as the non-diagnostic signal *u* does not provide any relevant information for the decision maker. Combining the above two equations yields the following.

$$\pi_{1k} = \left(\frac{\theta}{1-\theta}\right)^{z_t} \pi_0$$

Taking logs now yields:

$$\ln\, \pi_{1k} - \ln \pi_0 = z_t \cdot \left(\frac{\theta}{1-\theta}\right)$$

Finally, first differencing the above equation yields,

$$\Delta \ln\, \pi_{1k} = \Delta z_t \cdot \left(\frac{\theta}{1-\theta}\right)$$

where $\Delta z_t \in \{-1,0,1\}$. The first differenced equation demonstrates that – in absolute terms – a Bayesian agents updates, in log odds terms, at a constant of $\left(\frac{\theta}{1-\theta}\right)$.

Deriving the Regression Equation

Equation (1.5) is obtained by taking the natural logarithm of equation (1.3):

$$\ln\frac{\pi(G|S)}{\pi(B|S)}=\ln\left(\left[\frac{p(S|G)}{p(S|B)}\right]^{c_0+I\{u|\ desirable\}\cdot c_1+I\{u|\ undesirable\}\cdot c_2}\left[\frac{p(G)}{p(B)}\right]^{d}\right)$$

$$\Leftrightarrow \ln\frac{\pi(G|S)}{\pi(B|S)}$$

$$=(c_0+I\{u|\ desirable\}\cdot c_1+I\{u|\ undesirable\}\ \cdot\ c_2)\ \cdot\ln\frac{p(S|G)}{p(S|B)}+d\ \cdot\ln\frac{p(G)}{p(B)}$$

Next, that an agents' prior beliefs about the objective state of the world $\frac{p(G)}{p(B)}$ is equal to the agents' posterior belief from last period $\frac{\pi(G|s_1,\dots,s_{t-1})}{\pi(B|s_1,\dots,s_{t-1})}$.

$$\Leftrightarrow \ln\frac{\pi(G|S)}{\pi(B|S)}\ =(c_0+I\{u|\ desirable\}\cdot c_1+I\{u|\ undesirable\}\ \cdot\ c_2)\ \cdot\ln\frac{p(S|G)}{p(S|B)}$$

$$+d\cdot\ln\ \frac{\pi(G|s_1,\dots,s_{t-1})}{\pi(B|s_1,\dots,s_{t-1})}$$

Additionally, we accommodate the fact that a Bayesian agent updates his beliefs, in log odds terms, at a constant of $\Delta z_t\cdot\left(\frac{\theta}{1-\theta}\right)$. To do so, we follow Charness and Dave (2017) and Benjamin (2019), and replace ln $\frac{p(S|G)}{p(S|B)}$ with a dummy $D_{informative}$ taking the value 1 if the *t*th signal is *g*, 0 if the *t*th signal is *u*, and -1 if the *t*th signal is *b*. This alternate specification is equivalent, but the coefficient c_0 needs to be interpreted relative to $\left(\frac{\theta}{1-\theta}\right)$ instead of 1. To test whether individuals update their prior beliefs in response to non-diagnostic signals (i.e. when the *t*th signal is *u*), we additionally add two dummies, $D_{informative|\ desirable}$ and $D_{informative|undesirable}$ which equal 1 if the *t*th signal is u and if the signal is either in the positive, or in the negative domain, respectively:

$$\Leftrightarrow \ln\frac{\pi(G|S)}{\pi(B|S)}\ =c_0\cdot D_{informative}+c_1\cdot D_{informative|\ desirable}+c_2\cdot\ D_{informative|undesirable}$$

$$+\ d\cdot\ln\ \frac{\pi(G|s_1,\dots,s_{t-1})}{\pi(B|s_1,\dots,s_{t-1})}$$

Finally, note that the natural logarithm of subject's *i* odds ratio, based on her stated probability $P_{it}(G|s_1,\dots,s_t)$ that the asset is paying dividends from the good state is:

$$\lambda_{it} = \ln\left(\frac{\lambda(G|s_1, \dots, s_t)}{\lambda(B|s_1, \dots, s_t)}\right) = \ln\left(\frac{P_{it}(G|s_1, \dots, s_t)}{1 - P_{it}(G|s_1, \dots, s_t)}\right)$$

which may differ from the objective Bayesian probability. Incorporating this into the above equation, the final regression equation that we seek to estimate is as follows:

$$\Rightarrow \ln\left(\frac{\lambda(G|s_1, \dots, s_t)}{\lambda(B|s_1, \dots, s_t)}\right) = \widehat{\beta_1} \cdot D_{informative} + \widehat{\beta_2} \cdot \ln\frac{\pi(G|s_1, \dots, s_{t-1})}{\pi(B|s_1, \dots, s_{t-1})}$$

$$+\widehat{\beta_3} \cdot D_{informative|\ desirable} + \widehat{\beta_4} \cdot\ D_{informative|undesirable} + \varepsilon_{i,t}$$