Games and Economic Behavior 109 (2018) 21–39

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Belief updating and the demand for information

Sandro Ambuehl [a,*], Shengwu Li [b,*]

[a] Department of Management UTSC and Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada

[b] Society of Fellows, Harvard University, 78 Mt. Auburn Street, Cambridge, MA, 02138, USA

ARTICLE INFO

Article history:
Received 13 March 2016
Available online 8 December 2017

JEL classification:
C91
D01
D03
D83

Keywords:
Demand for information
Belief updating
Responsiveness to information
Probability weighting
Experimental economics

ABSTRACT

How do individuals value noisy information that guides economic decisions? In our laboratory experiment, we find that individuals underreact to increasing the informativeness of a signal, thus undervalue high-quality information, and that they disproportionately prefer information that may yield certainty. Both biases appear to be mainly due to non-standard belief updating. We find that individuals differ consistently in their *responsiveness to information* – the extent that their beliefs move upon observing signals. Individual parameters of responsiveness to information have explanatory power in two distinct choice environments and are unrelated to proxies for mathematical aptitude.

1. Introduction

Modern economies increasingly trade not just in physical goods, but in information. Economic agents consider whether to purchase medical tests, consulting expertise, or financial forecasts. Entire industries are premised on the sale of useful but noisy information.

In this paper we use a laboratory experiment to investigate whether, and why, the demand for information systematically deviates from the predictions of the standard rational agent model. In contrast to the existing literature on the demand for information, information in our setting is instrumentally useful but concerns a state that is of no intrinsic relevance to subjects, and strategic considerations play no role.

The purchase of information is commonly conceptualized as the decision to partake in a two-stage lottery. In the first stage, the agent updates her beliefs upon observing the realization of an informative signal. In the second stage, she chooses an action, upon which the state of the world is revealed and payoffs are realized.[1] A standard rational agent will update beliefs according to Bayes' rule, and make choices according to expected utility theory. For any agent who correctly anticipates his own belief updating, this suggests two mechanisms through which empirical behavior could deviate from the predictions of the standard model.

* Corresponding authors.
E-mail address: sandro.ambuehl@utoronto.ca (S. Ambuehl).

[1] This conceptualization of the demand for information has been chosen by a wide range of theoretical papers, including, for instance, Azrieli and Lehrer (2008), Athey and Levin (2001), Cabrales et al. (2013a), and Cabrales et al. (2013b).

https://doi.org/10.1016/j.geb.2017.11.009
0899-8256/

First, the demand for information may be affected by systematic deviations from expected utility theory in compound lotteries. These have been documented and modeled, for instance, by Bernasconi and Loomes (1992), Ergin and Gul (2009), Halevy (2007), Halevy and Feltkamp (2005), Segal (1987), Seo (2009), and Yates and Zukowski (1976). Second, the demand for information may be influenced by biases in belief updating (reviewed in Camerer, 1995). Our design allows us to estimate the effect of the latter channel on the demand for information, and to identify three biases in belief updating and the demand for information.[2]

In our experiment, subjects reveal how valuable they think it is to first observe an informative signal about a binary state of the world and then make a guess about it. At the very end of the experiment, a correct guess may be rewarded with a prize. Since our elicitation is in terms of probability units of winning the prize, our results are independent of the shape of subjects' utility function for money. Discriminating between channels is possible as a second stage asks subjects to reveal posterior beliefs for each of these signals. A third stage presents subjects with a conditionally i.i.d. sequence of signal realizations that depends on the state of the world. It measures how much information subjects require until they first prefer betting on the state of the world to an outside option.

We identify two biases in the demand for information that depend on the properties of *information structures*. Subjects value signals of varying informativeness as though they were more alike, compared to the predictions of the standard model (the *compression effect*). Additionally, subjects disproportionately prefer information structures that may perfectly reveal the state of the world (the *certainty effect*). We also test how valuations change if we increase the precision of the signal in one state of the world and decreasing it in another, such that the utility of a standard rational agent is unchanged. A class of theories of non-standard risk preferences predict that this will make an information structure less attractive to subjects (Segal, 1987; Seo, 2009). Our data do not support this hypothesis.

We identify a third bias that is *individual-specific*. Subjects in our experiment are heterogeneous in *responsiveness to information*, the extent to which their beliefs move upon observing information. Importantly, this heterogeneity is *consistent* within individuals. We measure this individual tendency in one task, and find that it is significantly correlated with subjects' behavior in two additional tasks. Moreover, while it is related to cognitive style, it is uncorrelated with proxies for mathematical aptitude. In particular, neither the direction nor the absolute size of this bias is correlated with whether a subject knows Bayes' law, has taken a statistics class, or has a STEM major.

Following Grether (1980) and Holt and Smith (2009), we propose a one-parameter model of responsiveness to information and estimate each subjects' parameter using data only from the latter part of our experiment. This procedure collapses all the data about each subject's belief updating to a single parameter, yet it explains 80% of the individual variation in the demand for information that can be explained with a model that uses each individual's beliefs data in a maximally flexible way. Our individual-level estimates of responsiveness to information are also significantly related to subjects' behavior in our environment that gradually uncovers information and reveals how quickly subjects' beliefs about the state of the world cross a given threshold.[3]

Conditional on the estimated parameters, the responsiveness model also generates the compression effect and the certainty effect. This is perhaps surprising, since the model is set up to capture a single individual-specific bias, while these two effects pertain to properties of the information structures. Both of these are consequences of the fact that our average subject is less responsive to information than a Bayesian, but correctly interprets perfectly informative and perfectly uninformative signal realizations. Formally, the model of responsiveness to information is equivalent to applying a prospect theory probability-weighting function to the Bayesian posterior. The information structure-specific effects we observe are implications of the inverse S-shape of that function.

Many of our results could in principle be due to either non-Bayesian belief updating, or due to non-standard risk preferences. We argue that the latter are unlikely to explain the entirety of our data, for three reasons. First, by design, our results cannot be due to the shape of subjects' utility function for money. Each choice in our experiment elicits, for some event X, the value of k such that the subject is indifferent between *receiving \$35 with probability k and \$0 otherwise*, and *receiving \$35 if event X occurs and \$0 otherwise*.[4] For any expected utility maximizer, the value of k that leaves such an agent indifferent does not depend on his degree of risk aversion (Roth and Malouf, 1979).

Second, also by design, they cannot be due ambiguity aversion, at least if interpreted in the standard fashion. In our experiment, all probabilities are objectively given. Subjects have induced priors about the state space, and all information structures are explicitly given, so that the probability of any payoff-relevant event can be calculated, via a simple application of Bayes' rule. In contrast, ambiguity aversion models are usually interpreted to apply under Knightian uncertainty, where objective probabilities are not applicable (Knight, 1921; Ellsberg, 1961; Baillon and Bleichrodt, 2015).

---

[2] This exercise, however, does not decompose behavior into beliefs and *preferences*. In the subjective expected utility framework, preferences are basic, and beliefs are just part of the representation (Savage, 1954). To be scrupulous, when we say that a subject deviates from Bayesian belief updating, we mean "in the representation that rationalizes this choice, the beliefs differ from the objective Bayes posteriors."

[3] We emphasize that heterogeneous responsiveness to information is not observationally equivalent to heterogeneous risk aversion. To see this, consider an agent who chooses between a safe option and a risky option, and who purchases information that might affect the optimal decision. If, in the absence of information, the agent would choose the risky option, then his willingness to pay for information is *increasing* in risk aversion. If, in the absence of information, the agent would choose the safe option, then his willingness to pay for information is *decreasing* in risk aversion. By contrast, his willingness to pay for information is *increasing* in responsiveness in both of these cases. (Appendix C formalizes this argument.)

[4] This mechanism was suggested in Allen (1987), Grether (1992), and Karni (2009), Schlag and van der Weele (2013), and has been used, amongst others, by Hoelzl and Rustichini (2005) and Moebius et al. (2013).

Third, there are other accounts of decision making under risk that are not excluded by our design choices. These do not account for the *compression effect* in the valuation of information structures, however, as this effect vanishes once we account for elicited belief updating data. The extent of the *boundary* effect is affected by controlling for beliefs data, but some of it remains even then. This suggests that it is partially rooted in risk preferences.[5]

Our study contributes to the literature in several ways. While responsiveness to information relates to both base-rate neglect (Kahneman and Tversky, 1972) and conservatism (see Peterson and Beach, 1967 for a review), we show that it is *heterogeneous* across subjects, *stable* within subjects, and has explanatory power across three choice environments.[6] We explicitly relate biases in belief updating to prospect theory probability weighting (as reviewed, for instance, in Fehr-Duda and Epper, 2012).

A handful of papers have studied individual heterogeneity in belief updating. Most closely related is the successor to our paper by Buser et al. (2016) who also find, in a setting with ego-relevant beliefs, that responsiveness to feedback is a personal trait that predicts selection into competitive environments. (They do not relate the trait to information demand, and do not consider variation in information structures.) Augenblick and Rabin (2015) show that if a forecaster's belief stream for a given geopolitical event is relatively volatile, then the same forecaster's belief stream for other events is also relatively volatile. El-Gamal and Grether (1995) fit a finite mixture model to binary-choice data, and find that the most frequently used updating rules are Bayes, base-rate neglect (which roughly corresponds to excessive responsiveness), and conservatism (which is equivalent to insufficient responsiveness). These authors do not study individual consistency across different kinds of updating tasks.[7] Our paper complements Antoniou et al. (2017) who document individual heterogeneity in how strongly individuals react to the salience of information, keeping constant its diagnosticity.

Within the empirical literature on the demand for information, our study contributes to the literature by investigating biases pertaining to different information structures (characterized by their state-dependent signal accuracies), and by relating them to a stable individual trait that characterizes belief updating. These results distinguish it from Hoffman (2016) who studies internet-domain-name traders in a lab-in-the-field experiment. He finds that the willingness to pay (WTP) for information increases with a subject's uncertainty about the state of the world, increases with the accuracy of the signal, as predicted by the Bayesian benchmark, but does so to a lesser extent than predicted by theory. This is related to the compression effect we document in the demand for information.

A related literature documents that subjects do not always sufficiently account for the properties of the data generating process when making inferences (Massey and Wu, 2005; Fiedler and Juslin, 2006). Our finding that subjects' are willing to pay too much for low-quality information, and not willing to pay enough for high-quality information complements Falk and Zimmermann (2016) who find that in a context of *non*-instrumental information, subjects prefer few realizations of high-quality information to many realizations of low-quality information.[8]

We know of no further empirical work on the demand for information in a non-strategic setting in which information is instrumentally useful but concerns a state that is of no intrinsic relevance to subjects, and strategic considerations play no role. Outside of our domain, Eliaz and Schotter (2010) show that subjects are willing to pay for information that will not change their decision, if it decreases their uncertainty about having taken the right action, and Bastardi and Shafir (1998) find that once subjects decide to purchase information, they base their decisions on it, even if they would not do so if they were given the information for free. Masatlioglu et al. (2015) study how the demand for non-instrumental information varies with the properties of the available information structures. Burks et al. (2013), Eil and Rao (2011), and Moebius et al. (2013) find that when beliefs are ego-relevant, WTP for information is larger the more confident the subject is that information will be desirable. Kocher et al. (2014), Falk and Zimmermann (2016), and Zimmermann (2014) study preferences over the dynamic arrival of information. The first of these papers finds that subjects prefer later resolution of uncertainty about a lottery ticket, the second finds that they prefer later resolution about whether an unpleasant event will occur if they are not forced to focus on that event, and the third that subjects display no preference over clumped vs. piecewise information regarding a lottery ticket. The demand for dynamic information is also studied in the literature on sequential search (for instance Brown et al., 2011; Caplin et al., 2011, and Gabaix et al., 2006) and in the literature on explore–exploit dilemmas (see Erev and Haruvy, 2015 for a recent review). Finally, endogenous information acquisition has been studied in a wide range of strategic settings. For example, Kubler and Weizsacker (2004) focus on social learning and informational cascades,

[5] We do not have data, however, to allow us to make claims about the extent of *individual-level* heterogeneity that is due to non-standard belief updating as opposed to non-standard risk preferences. Therefore, we make no claims about the extent of total individual-level variation our measure of responsiveness captures; but merely that it captures the bulk of that part of the individual-level variation *that can maximally be explained by non-standard belief updating*.

[6] The literature on belief updating is vast, and we do not attempt to review it here. See Camerer (1995) for a review. Benjamin et al. (2015) contains a recent comprehensive review of belief updating experiments that use binary states and binary signals.

[7] Individual heterogeneity in belief updating was observed (but not explicitly investigated) in Peterson et al. (1965). Moebius et al. (2013) also find heterogeneity in belief updating in an ego-relevant task, and show that beliefs are equally predictive of subsequent behavior separately for the more and less conservative half of subjects.

[8] Our work may appear related to a literature on decision making under ambiguity (Gilboa and Marinacci, 2013). We do not interpret our data in light of this literature, for two reasons. First, our experiment involves only objective uncertainty. Second, the literature of decision making under ambiguity has not yet provided a canonical model of belief updating. In contrast, there is a long tradition of analyzing belief updating with objective probabilities (e.g. Peterson and Beach, 1967), as well as on deviations from expected utility theory when objective probabilities are given (e.g. Kahneman and Tversky, 1979). Our analysis is closely related to these strands of literature.

Szkup and Trevino (2015) on global games, Gretschko and Rajko (2015) on auctions, and Sunder (1992) and Copeland and Friedman (1987) on experimental asset markets.

The remainder of this paper proceeds as follows: Section 2 presents the experimental design and our hypotheses. Section 3 analyzes behavior as a function of the information structure, documents the compression and certainty effects, and argues that they are likely due to belief updating biases rather than due to non-standard risk preferences. Section 4 studies individual-level consistency, presents our results regarding responsiveness to information, and links them to the information-structure specific biases. Finally, section 5 discusses potential applications of our findings and concludes.

2. Experimental design

We present a brief overview of the design before discussing specifics and hypotheses below.

We consider a setting in which information is valuable because it informs a subsequent choice. We focus on the *prediction game*, which is arguably the simplest nontrivial such setting. The state of the world s is either 0 or 1. An agent with some prior belief observes an informative signal about the state, and guesses the state of the world. If (and only if) his guess is correct, he receives a prize k.

We are interested in how agents value such an informative signal, and how this depends on its properties. Hence, we elicit the probability v such that the agent is indifferent between *receiving the prize with probability* v and *playing the prediction game*. This is the *information valuation task*.

If these valuations differ from those a Bayesian expected utility maximizer would have, we would like to pinpoint the source of the deviation. The experiment thus separately examines each of the steps a rational agent would make to decide about her valuation. These involve calculating the posterior belief she would obtain and the action she would take, for each possible signal realization, on the one hand, and calculating the likelihood of observing each signal, on the other. To arrive at the valuation v_I, the Bayesian then simply aggregates these probabilities. Accordingly, we elicit subjective posterior beliefs (the *belief updating task*), as well as subjective probabilities of observing each signal (the *probability assessment task*). We elicit these elements only *after* eliciting valuations of information structures, so we do not nudge the subjects towards thinking about information valuation in a specific fashion.

Finally, if we find that empirical belief updating behavior is non-Bayesian, we would like to study the extent to which these deviations are systematic and stable within individuals. The experiment ends with the *gradual information task* which allows us to answer this question by testing how closely deviations in one task are associated with deviations in another task.

We now discuss our setup in detail, outline our hypotheses, and describe how we implement the experiment.

2.1. Formal setup and hypotheses

Setup We use $s \in \{0,1\}$ to denote the state of the world, about which subjects have an objective prior belief $P(s=1)=\frac{1}{2}$. We let $\sigma \in \{0,1\}$ denote is the realization of a stochastic signal the subject may observe. For each such signal, subjects are aware of the pair of state-dependent probabilities $P_I(\sigma=1|s=1)$ and $P_I(\sigma=0|s=0)$. We refer to such a pair $I=\big(P_I(\sigma=1|s=1), P_I(\sigma=0|s=0)\big)$ as an *information structure*. Notice that, even though both states are equally likely, both signals need not be equally likely.

We let $v_{I,i}$ denote the probability that renders subject i indifferent between playing the prediction game with information structure I, on the one hand, and receiving the prize with objective probability $v_{I,i}$, on the other. We refer to $v_{I,i}$ as agent i's *valuation* of information structure I. Because there are only two possible levels of payoff a subject can obtain (she wins the prize or does not), this measure of valuation is independent of the curvature of a subject's utility function for money.

Observe that *subjective* posterior beliefs $P_{I,i}(s|\sigma)$ and signal probability assessments $P_{I,i}(\sigma)$ are sufficient to obtain the valuation of information structure I for any agent who reduces compound lotteries, has preferences that depend only on final outcomes, and strictly prefers obtaining the prize k to not obtaining it — even if he is not Bayesian. Given a strategy $g(\sigma)$ that maps a signal realization into a guess about the state of the world, i's valuation of information structure I is given by

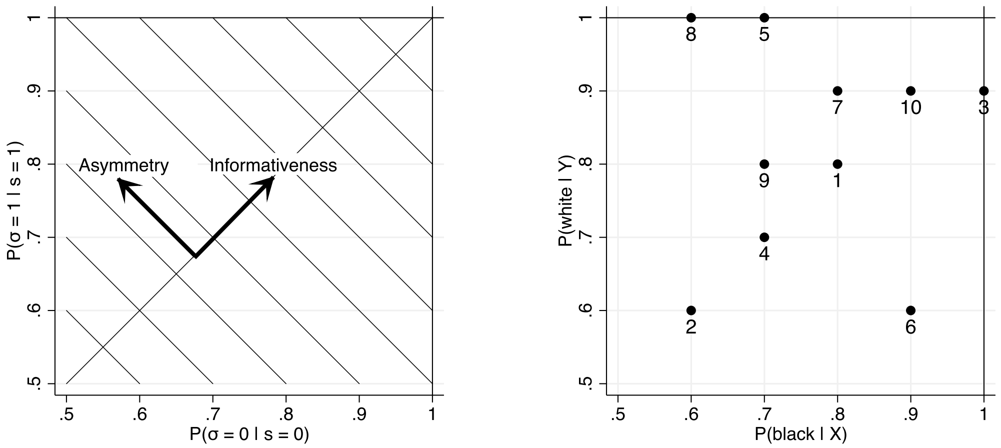
$$v_{I,i} = P_{I,i}\big(g(\sigma)|\sigma=1\big)P_{I,i}(\sigma=1) + P_{I,i}\big(g(\sigma)|\sigma=0\big)P_{I,i}(\sigma=0). \tag{1}$$

Our design imposes $P_I(\sigma=1|s=1) > 0.5$ and $P_I(\sigma=0|s=0) > 0.5$, so that the optimal guess is $g(1)=1$ and $g(0)=0$. Then the above equation reduces to:

$$v_{I,i} = P_{I,i}\big(s=1|\sigma=1\big)P_{I,i}(\sigma=1) + P_{I,i}\big(s=0|\sigma=0\big)P_{I,i}(\sigma=0). \tag{2}$$

Our experiment elicits each constituent part of equation (2). The *information valuation task* elicits $v_{I,i}$, and the *belief updating* and *probability assessment* tasks elicit $P_{I,i}(s|\sigma)$ and $P_{I,i}(\sigma)$, respectively, for each possible signal realization $\sigma \in \{0,1\}$.

Finally, in the *gradual information task*, we present subjects with a different kind of belief updating problem to further test whether subjects' responsiveness to information has explanatory power in other decision environments. Subjects are



(A) (B)

Fig. 1. Panel (A) displays the set $[0.5, 1] \times [0.5, 1]$ of information structures and the level curves of v_I^{Bayes}. We call an information structure I' that lies to the northeast of I more *informative* than I. We call an information structure I' that lies to the northwest of I more *asymmetric* than I. Panel (B) displays the set of information structures we used in the experiment and the order in which subjects went through them. (Approximately half of the subjects proceeded through the information structures in reverse order.) The horizontal axis measures the proportion of black balls in box X and the vertical axis measures the proportion of white balls in box Y.

shown a sequence of conditionally independent signals that are informative about the state of the world, and decide, after each signal, whether they prefer to take part in a lottery with an explicitly given winning probability, or whether they prefer to bet on the state of the world.

Hypotheses In our setting, the space of all information structures is the unit square $[0, 1] \times [0, 1]$, which we can display graphically. We focus on the subspace $[\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$ on which following the signal is the optimal guessing strategy, $g(\sigma) = \sigma$.

How would a Bayesian agent value the prediction game with an information structure I? By replacing $P_{I,i}(s|\sigma)$ and $P_{I,i}(\sigma)$ in equation (2) by their objective counterparts, we derive

$$v_I^{Bayes} = \frac{P_I(\sigma = 1|s = 1) + P_I(\sigma = 0|s = 0)}{2} \tag{3}$$

Fig. 1(A) plots the level curves of v_I^{Bayes} over the set of information structures under consideration. We call v_I^{Bayes} the *informativeness* of information structure I; it is the probability that the signal matches the state of the world, and it increases to the northeast of the figure.

We study information structure-specific biases by answering the following questions. 1. Does subjects' valuation of information increase one-for-one with the Bayesian valuation as we move an information structure along the diagonal, or do subjects under- or overreact to such a change? 2. Do subjects have disproportionately high or low valuations for information structures on the northern or eastern boundary of Fig. 1(A), which provide certainty for at least one of the signal realizations? 3. Are the empirical indifference curves indeed linear in the interior of $[\frac{1}{2}, 1] \times [\frac{1}{2}, 1]$, or do subjects have preferences for (or against) asymmetric information structures, which would cause the indifference curves in Fig. 1(A) to be strictly concave (convex)? 4. Are deviations from the theoretical predictions due to non-Bayesian belief updating, or are they caused by non-standard preferences over two-stage lotteries?

We then focus on individual-specific biases and investigate whether an individual's behavior in simple belief updating tasks is predictive of her behavior in other choice problems. We fit a single-parameter model using only each subject's behavior in belief updating tasks, and employ that to predict her behavior in information valuation task and in the gradual information task. Finally, we tie this model to the information structure-specific biases we document in the first part of the analysis.

2.2. Implementation

In most parts of the experiment, subjects face computerized pairs of boxes filled with black and white balls, as in Fig. 2. Box X contains 10 balls, the majority of which are black. Box Y contains 20 balls, the majority of which are white. In many parts of the experiment, the computer first randomly and secretly selects a box, and then draws a ball from the selected box for the subject to see. The box selected corresponds to the state of the world s, while the color of the ball drawn

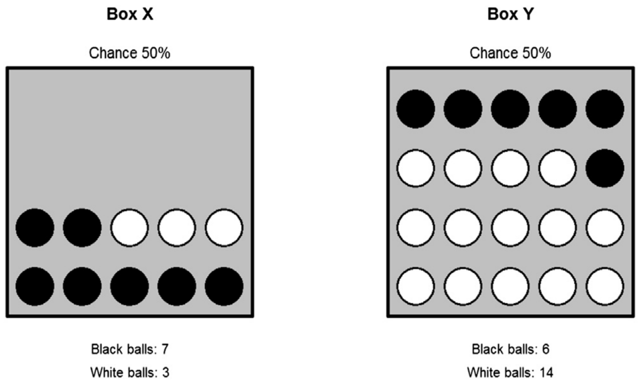


Fig. 2. Screenshot from the experimental interface.

corresponds to the signal realization σ. We chose different ball totals in Box X and Box Y in order to ensure that naïve ball-counting heuristics are not observationally equivalent to Bayesianism in the *Information Valuation Task*.[9]

We ensure incentive compatibility by paying subjects for a single decision randomly selected from the entire experiment. Subjects never receive feedback about their decisions until the very end of the experiment. Each subject participates in each of the following five parts of the experiment, in the order listed here.

Part 1: prediction game Subjects play six rounds of the prediction game with different information structures so that they can familiarize themselves with the game that they are subsequently asked to value.[10] In each round, the computer randomly and secretly selects a box, each with 50% probability, and shows the subject a ball drawn at random from that box. The subject then guesses which box was selected, receiving \$35 (\$0) for a correct (incorrect) guess. Subjects play the prediction game for six different information structures.

Part 2: information valuation task This part elicits subjects' valuations of information structures $v_{I,i}$. Subjects are presented with the 10 pairs of boxes depicted in Fig. 1(B). We chose them so as to easily test our hypotheses.[11] For each pair of boxes, subjects fill in a multiple choice list, choosing between pairs of options: "Play the Prediction Game" or "Win \$35 with chance p", for p between 1% and 100%, in 3% intervals. The value p at which a subject switches from preferring the former to the latter option reveals her valuation $v_{I,i}$ of the Prediction Game with those pairs of boxes. Since each line counts as a separate decision, one of which might be randomly drawn for being carried out, truthful revelation is strictly optimal. We constrain subjects to have at most one switching point.[12] Half the subjects, chosen at random, evaluate the ten information structures in the order indicated in Fig. 1(B), the remaining subjects evaluate them in reverse order.

If at the end of the experiment a decision from this stage is randomly chosen for implementation, a subject who has chosen to play the prediction game first sees a ball drawn from one of the two boxes, and then decides which box to bet on. The subject receives the prize if, and only if, her guess coincides with the box from which the ball was drawn.

Part 3: belief updating task We use the strategy method to elicit posterior beliefs $P_{I,i}(s=0|\sigma=0)$ and $P_{I,i}(s=1|\sigma=1)$ for each of the ten information structures in Fig. 1 and for each possible signal realization. Subjects fill in a choice list to be implemented if the drawn ball is black, and a choice list to be implemented if the drawn ball is white. These consists of the lines "Receive \$35 if the Selected Box is Box X" and "Receive \$35 with chance p", for p between 1% and 100%, in 3% intervals. Subjects know that if a round from this is drawn for payment, a ball will be drawn from the selected box at random, and their decision on a randomly chosen line of the corresponding choice list will be implemented.[13] Subjects make their choices in the same order as in part 2.

Part 4: eliciting signal probabilities We elicit the subjects' signal probability assessments $P_{I,i}(\sigma)$ for each of the 10 information structures using multiple choice lists. The choice on each line is between an exogenous probability of winning \$35, and

[9] If there were 10 balls in each box, then the result of the following heuristic would coincide with the Bayesian valuation: Count the number of black balls in box X, count the number of white balls in box Y, sum them up and divide by 20. Appendix B.2 shows that this difference in urn sizes does not affect our results.

[10] We draw signals independently and randomly for each subject, conditional on state, according to the current information structure. The information structures used are $\{(0.5, 0.9), (0.7, 1), (0.6, 0.6), (1, 0.55), (0.6, 0.85), (0.8, 0.6)\}$. We use six information structures so that each subject is quite likely to see both black and white draws, but also so that this part of the experiment is concluded quickly.

[11] The choice of information structures also reflects power calculations calibrated to the test by Abrevaya and Jiang (2005). After obtaining the data, we discovered that this test is subject to substantive bias with small sample sizes, and hence we do not apply it.

[12] Andersen et al. (2006) finds that enforcing a single switching point in a multiple decision list has no systematic effect on subject responses.

[13] Holt and Smith (2016) show that this mechanism elicits a higher frequency of Bayesian responses than alternative belief elicitation mechanisms.

Table 1
Choice problems in the gradual information task.

Block	Information structure	q
1	$(0.7, 0.7)$	0.8
2	$(0.8, 0.8)$	0.9
3	$(0.7, 0.7)$	0.8
4	$(0.6, 0.6)$	0.75

winning the same amount if a ball of a specified color will be drawn from the box that was to be selected at random from the given pair. The computer randomly determines whether a subject is asked to assess the likelihood of a black or a white ball. Subjects make their choices in the same order as in part 2.

Part 5: gradual information task This task is played in four *blocks*. In each of them, subjects see a pair of boxes, each containing some combination of 20 balls (black or white). At the start of the block, the computer randomly selects a box, each with 50% probability, and shows the subject a sequence of 12 draws *with replacement* from the Selected Box. After *each* draw, subjects indicated whether they would rather (i) receive \$35 if Box X is the Selected Box (ii) receive \$35 if Box Y is the Selected Box, or (iii) receive \$35 with probability q for some fixed success probability q. No choice prematurely ends the sequence of draws, or precludes any future choices. Subjects face, in order, the information structures and values of q in Table 1.[14]

Questionnaire The final part of the experiment is a questionnaire that includes demographic variables, psychological measures, and the questions from the Cognitive Reflection Test (CRT, Frederick, 2005).

Procedures We conducted the experiment at the Stanford Economics Research Laboratory (SERL) using z-tree (Fischbacher, 2007). Undergraduate and graduate students were recruited from the SERL subject pool database using standard email recruiting procedures. We excluded graduate and undergraduate students in psychology as well as graduate students in economics from participation. A total of 143 individuals participated in 9 sessions of the experiment, in addition to 13 individuals for a pilot session.[15]

All sessions were run in May 2013, and each experimental session took between 2 and 2.5 hours to complete, including payment. We distributed and read aloud the instructions directly before each part and informed subjects that their choices in any part would not affect their earnings from any other part. We provided subjects with pen and paper, and neither encouraged nor discouraged their use. The same two experimenters were present in each session.

Each subject received a \$5 show-up payment, an additional \$15 for completing the experiment, and played for a prize of \$35. In addition, a subject could earn \$1 per CRT question answered correctly.

Conventions We code the data such that for every information structure I, $P_I(\sigma = 1|s = 1) \geq P_I(\sigma = 0|s = 0)$. Hence, $\sigma = 1$ is the more likely and thus less informative signal realization.

As Fig. 3, panel (C) indicates, the information structures in this experiment can be divided into three groups: Information structures S1, S2, S3, and S4 are evenly spaced symmetric information structures. Information structures A1, A2, and A3 are asymmetric information structures that are not on the boundary. Information structures B1, B2, and B3 are on the boundary. They afford certainty for one realization of the signal.

Since we have used multiple choice lists in with increments of 3 percentage points, our data are interval coded. Throughout, we use the midpoint of the interval for analysis.

Subject understanding Before conducting data analysis, we test whether subjects understood that following the signal is the optimal strategy in the prediction game. In part 1 of the experiment, 83% of our subjects made the correct prediction for all six information structures and just 4 out of 143 subjects (2.8%) made more than 2 mistakes.[16] We therefore proceed on the assumption that subjects' valuation of information structures is based on anticipation of this optimal strategy.

3. Information structure-specific biases

How well does the standard model predict average valuations of information structures? Fig. 3, panel (A), plots the mean difference (across subjects) between the elicited and the Bayesian valuation for each information structure. Information structures are arranged by increasing informativeness. Perhaps the most salient aspect of Fig. 3, panel (A), is the fact that

[14] These parameter choices imply that a Bayesian decides to bet on a box as soon as he has seen two more balls of one rather than the other color in blocks 1 to 3, and three more balls in block 4.

[15] Two sessions presented subjects with the additional information structure $(0.5, 0.5)$. One observation for one subject (out of 1430 total observations) was not recorded due to a database error.

[16] Moreover, 43.8% of all mistakes occur for the information structure $I = (0.6, 0.6)$, the information structure for which such errors are cheapest.

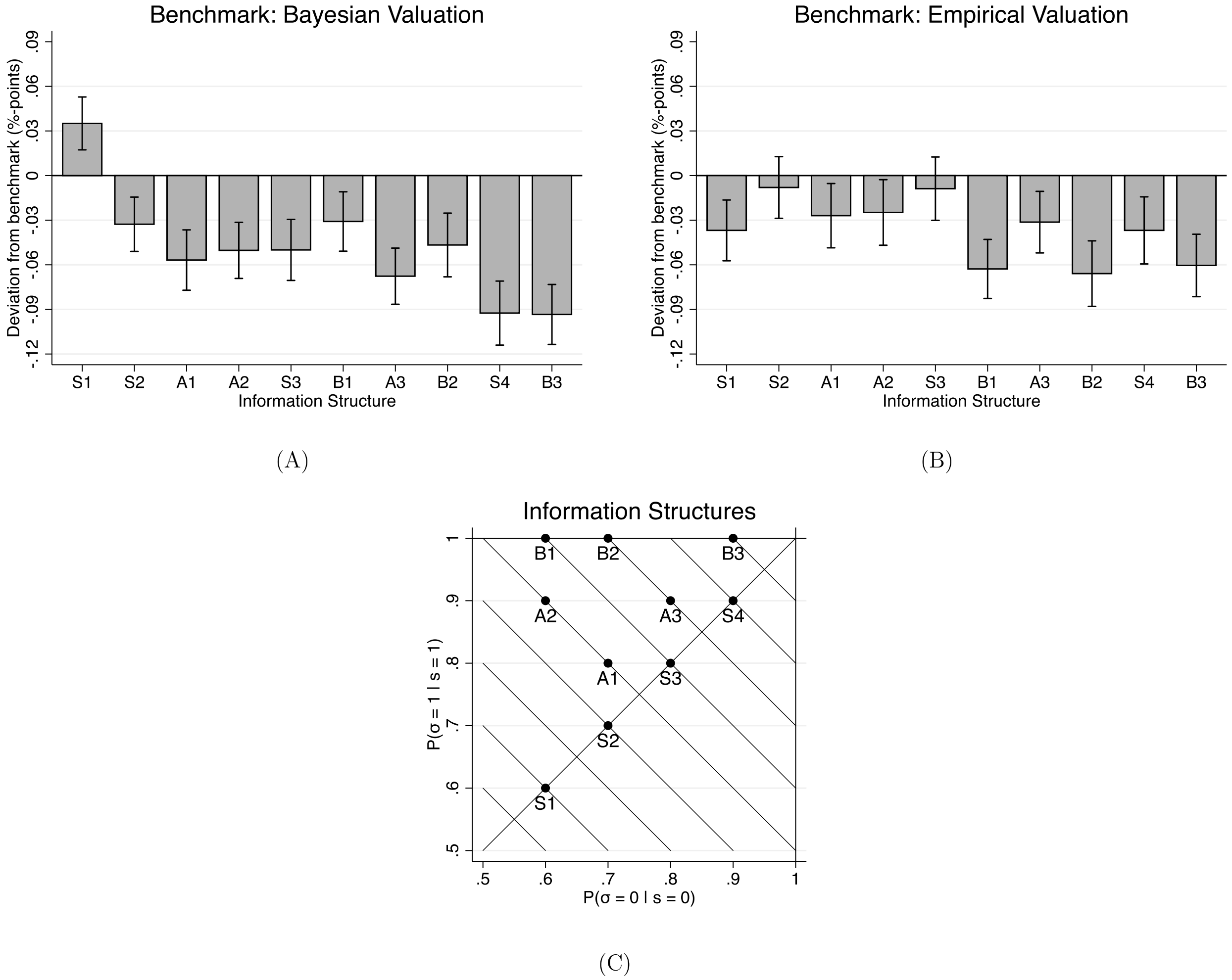


Fig. 3. Panel (A) depicts the mean deviation of the valuation of an information structure from the Bayesian benchmark, Panel (B) depicts the mean deviation from the subjective beliefs benchmark. Information structures are arranged in order of increasing informativeness. Standard errors are clustered by subject. For reference, panel (C) displays the information structures. Symmetric information structures are labeled $S1$–$S4$. Asymmetric information structures are $A1$–$A3$. Boundary information structures are $B1$–$B3$. For readability, we code the information structures such that $\sigma = 1$ is the weakly more frequent signal for each information structure. In the experiment, there was variation across information structures whether a black or a white ball was the more frequent (and hence less informative) signal realization; see Fig. 1.

subjects undervalue all but one information structure, and undervalue them to a greater extent the more informative they become. In addition, boundary information structures seem to be valued more highly than equally informative information structures in the interior. Meanwhile, there is no indication of a bias relating specifically to asymmetric (but interior) information structures.

Formally, we estimate the model

$$v_{I,i} = \beta_0 + \beta_1 v_I^{Bayes} + \beta_2 a_I + \beta_3 b_I + \epsilon_{I,i} \tag{4}$$

Here, $a_I = \frac{P(\sigma=1|s=1)-P(\sigma=0|s=0)}{2}$ is the *asymmetry* of an information structure I, measured by its off-diagonal distance from the 45-degree line.[17] b_I is a dummy variable that equals 1 if I is a boundary information structure, and 0 otherwise. The standard rational agent model requires that $\beta_1 = 1$ and $\beta_0 = \beta_2 = \beta_3 = 0$.

Results are reported in Column 1 of Table 2. β_1 is both significantly less than 1 and significantly larger than 0. Hence, while subjects' valuations increase with informativeness, they do not do so by a sufficient amount. We call this result the *compression effect*.[18] We also find that subjects have a significant preference for information structures where one signal

[17] Recall that we normalized the data such that $a_I \geq 0$ for every information structure I.
[18] Appendix B.1 shows that the result is robust to controlling for truncation at 1 and 0.5.

Table 2
Aggregate valuations of information structures. Reported significance levels for v_I^{Bayes} concern the hypothesis that this coefficient is 1. Standard errors clustered by subjects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Info structures	[S,A,B]	[S,A,B]	[S]	[S,A]	[S,B]	[S,B]	[S,A]
Dependent VAR	$v_{I,i}$	$v_{I,i}$	$v_{I,i}$	$v_{I,i}$	$v_{I,i}$	$v_{I,i}$	$v_{I,i}$
v_I^{Bayes}	0.603***	0.677***	0.594***	0.616***	0.655***	0.491***	0.618***
	(0.0427)	(0.0386)	(0.0450)	(0.0439)	(0.0406)	(0.0427)	(0.0437)
Asymmetry	−0.0237						−0.0475
	(0.0260)						(0.0295)
Boundary	0.0342***					0.0273***	
	(0.00968)					(0.00843)	
Constant	0.260***	0.208***	0.268***	0.248***	0.231***	0.271***	0.249***
	(0.0326)	(0.0305)	(0.0338)	(0.0326)	(0.0323)	(0.0330)	(0.0326)
Observations	1429	1429	571	1000	1000	1000	1000
R-squared	0.177	0.172	0.186	0.140	0.202	0.207	0.141

realization removes all uncertainty, on the order of 3.4 percentage points, and hence are subject to a *certainty effect*.[19] We do not find evidence for asymmetry effects.

Why do subjects undervalue rather than overvalue most information structures? The estimates in Column 1 imply that subjects correctly value information structures with informativeness 0.653 (s.e. 0.024), and overvalue (undervalue) those with lower (higher) informativeness.[20] Accordingly, subjects undervalue most information structures in our experiment because we included a disproportionate number of information structures with comparatively high informativeness (in order to better study asymmetry and boundary effects).

To be sure that our results are not driven by functional form mis-specification, we check that they are not changed when assessing cleanly comparable subsets of the data. The compression effect persists when we no longer include a_I and b_I in the regression (Column 2), when we consider only symmetric, evenly spaced information structures (Column 3), when we exclude all boundary information structures (Column 4), and when we exclude all asymmetric non-boundary information structures (Column 5). Certainty effects persist, when we only compare boundary information structures to symmetric information structures (Column 6), and their magnitude is not substantially changed. We do not find significant asymmetry effects, even when we compare only non-boundary asymmetric information structures with symmetric information structures (Column 7).[21]

3.1. Causes

There are at least two channels through which these biases in the demand for information could arise. On the one hand, they could be caused by subjects' anticipation of non-standard belief updating (the *beliefs channel*). On the other hand, subjects could exhibit non-standard information valuations because, even *conditional on their own beliefs*, they have behavioral traits, such as non-standard risk preferences, that affect how they value information, or because they fail to correctly forecast their own belief updating.

To isolate the beliefs channel, we use the data elicited in the Belief Updating Task to derive how a subject values an information structure if the preferences channel is absent, the subject correctly forecasts his belief updating, but belief updating is possibly non-Bayesian. We call this the *subjective-beliefs prediction* of the valuation, and define it by

$$v_{I,i}^{Beliefs} = P_{I,i}(s=1|\sigma=1)P_{I,i}(\sigma=1) + P_{I,i}(s=0|\sigma=0)P_{I,i}(\sigma=0) \tag{5}$$

To assess the impact of the remaining channels, Fig. 3, panel (B), compares subjects' valuations of an information structure to the subjective-beliefs prediction. The compression effect seems to disappear, as does the certainty effect. However, average valuations fall significantly short of the predicted valuations, implying that conditional on their own updated beliefs, subjects on average undervalue information.

We formally establish that non-standard belief updating rather than non-standard preferences or a failure to correctly forecast belief updating is the reason for the compression effect and the certainty effect. Briefly, our approach is the following. Equation (5) predicts how an expected utility maximizer should value an information structure if she correctly forecasts her own belief updating, which, however, may differ from the Bayesian benchmark. If non-standard valuations of information structures derive entirely from non-standard belief updating, then elicited valuations $v_{I,i}$ should not systematically

[19] This result is virtually unchanged when we exclude all individuals who expressed a posterior two or more categories below 100% in the cases where the Bayesian posterior is 1.

[20] We calculate this by solving $v = \beta_0 + \beta_1 v$ for v, and replacing the parameters β_0 and β_1 with the estimates from Column 1.

[21] To assess the magnitude of the estimated coefficient, note that the information structure $(0.5, 1)$ has the maximum possible asymmetry, which is 0.25. This information structure is predicted to be valued only 1.2 percentage points less highly than the equally informative, perfectly symmetric information structure $(0.75, 0.75)$.

Table 3
Relationship between beliefs and valuations. Reported significance levels for $v_I^{Beliefs}$ concern the hypothesis that this coefficient is 1. Standard errors clustered by subjects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

	(1)	(2)	(3)
Method	OLS	OLS	IV
Dependent var	$v_{I,i}$	$v_{I,i}^{Beliefs}$	$v_{I,i}$
$v_{I,i}^{Beliefs}$	0.471***		0.984
	(0.0775)		(0.0730)
$v_{I,i}^{Bayes}$		0.613***	
		(0.0280)	
Asymmetry	−0.0719	−0.0108	−0.0131
	(0.0279)	(0.0209)	(0.0335)
Boundary	0.0443***	0.0687***	−0.0336**
	(0.0134)	(0.0077)	(0.0141)
Constant	0.373***	0.276***	−0.0118
	(0.0612)	(0.0229)	(0.0540)
Observations	1429	1429	1429
R-squared	0.222	0.316	0.0524

differ from these predictions $v_{I,i}^{Beliefs}$. By contrast, if non-standard preferences play a substantial role (or if subjects fail to correctly anticipate their own belief updating when valuing information structures), we should detect systematic deviations between $v_{I,i}$ and $v_{I,i}^{Beliefs}$.

A naïve testing strategy would regress $v_{I,i}$ on $v_{I,i}^{Beliefs}$, as well as the characteristics of the information structures, a_I, and b_I. The results from this are reported in Column 1 of Table 3. Because beliefs, and therefore $v_{I,i}^{Beliefs}$, may be measured with error, however, attenuation bias could cause inconsistent estimates. Consequently, we instrument for the subjective-beliefs prediction of valuations with the Bayesian valuation.[22] The first-stage regression reported in Column 2 confirms that the instrument is relevant. Column 3 contains the coefficients of interest. The coefficient on $v_{I,i}^{Beliefs}$ is statistically indistinguishable from one ($p = 0.229$), which implies that the compression effect is indeed unrelated to non-standard preferences. The boundary dummy is slightly, but significantly, negative. Hence, the positive certainty effects in the previous section operate mainly through the beliefs channel; the remaining channels add a countervailing effect. Again, there is no evidence of asymmetry effects.

4. Individual-specific biases

In this section we study individual-specific biases, and tie them to the information-structure-specific biases. We find that deviations from the Bayesian benchmark are consistent within individuals: a subject whose beliefs move too much upon observing a signal from some information structure tends to update his beliefs too much for all information structures, and by a similar amount.

We employ a one-parameter model that captures much of the variation in behavior across subjects. We estimate, for each individual, their *responsiveness to information* using only data from the belief updating task, and show that it is significantly correlated with behavior in two different domains. First, it explains more than 80% of the non-standard valuation behavior that is attributable to non-standard belief updating. Second, it is significantly correlated with behavior in the Gradual Information Task. This suggests that responsiveness to information is an individual tendency that is stable across different choice environments. It is orthogonal to proxies of mathematical aptitude such as college major and self-reported knowledge of Bayes' law.

Given the estimated parameters, our model also generates the compression effect and the certainty effect that we documented in the preceding section (but predicts magnitudes that fall short of those we observe). This is not an obvious consequence, since these are information-structure-specific effects, and nothing in the responsiveness model directly relates to properties of information structures.

4.1. Responsiveness to information

For both the information valuation and belief updating tasks, we observe a subject's choices for 10 distinct information structures. If subjects have internally consistent *directional* biases, *within* each of these tasks, a subject's mean deviation from

[22] The identifying assumption is that the correct Bayesian beliefs affect subjects' valuations only *via* their subjective beliefs. This exclusion restriction would not hold if, for instance, subjects were subconsciously influenced by the correct Bayesian posteriors when evaluating information structures, but this influence was not reflected in their *ex post* beliefs. Note that this exercise also establishes that beliefs have a *causal* effect on the demand for information.

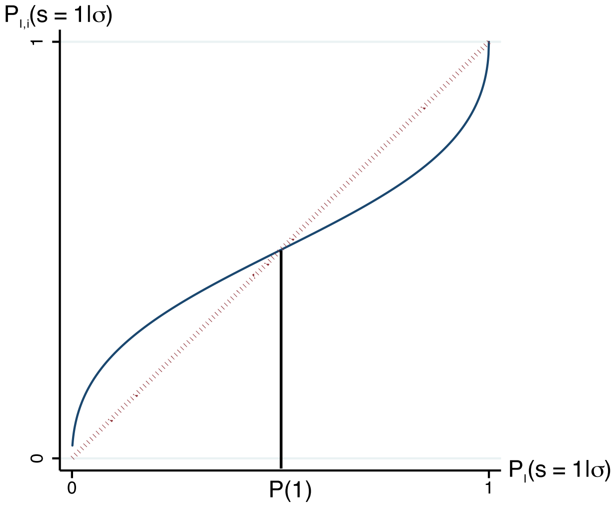


Fig. 4. Relation between the posterior $P_{I,i}(s = 1|\sigma)$ of a decision maker with responsiveness to information $\alpha_i = 0.5$ (plotted on the vertical axis) and the Bayesian posterior $P_I(s = 1|\sigma)$ (plotted on the horizontal axis) and prior $P_I(1)$.

the Bayesian benchmark on two of the information structures should be predictive of her mean deviation on the remaining eight.

Formally, for each information structure I we define subject i's deviation from the Bayesian benchmark by $\Delta v_{I,i} := v_{I,i} - v_I^{Bayes}$ in the information valuation task and by $\Delta P_{I,i}(s|\sigma) = \text{logit}(P_{I,i}(s|\sigma)) - \text{logit}(P_I(s|\sigma))$ in the belief updating task, where $\text{logit}(x) = \log(x/(1-x))$. We then calculate the extent to which subjects' mean deviation on two information structures correlates with the deviation on the remaining eight, for all possible such splits of the set of information structures, and average. These mean correlations are 0.773 (s.e. 0.035) in the case of $\Delta v_{I,i}$, 0.781 (s.e. 0.035) in the case of $\Delta P_{I,i}(s = 1|\sigma = 1)$, and 0.749 (0.075) in the case of $\Delta P_{I,i}(s = 0|\sigma = 0)$.[23] Since all of them are highly significantly positive, we conclude that subjects are internally consistent within tasks.[24]

This motivates our model of *responsiveness to information*, the extent to which a subject's beliefs move upon observing information. This model is adapted from Grether (1980) and Holt and Smith (2009).[25] It explains the individual consistency just documented, and allows us to relate behavior in different parts of the experiment to each other. To formally define responsiveness to information, note that a Bayesian updates beliefs according to the following formula with $\alpha_i = 1$:

$$\text{logit}(\underbrace{P_{I,i}(s = 1|\sigma)}_{\text{posterior}}) = \text{logit}(\underbrace{P(s = 1)}_{\text{prior}}) + \alpha_i \underbrace{\log\left(\frac{P_I(\sigma|s = 1)}{P_I(\sigma|s = 0)}\right)}_{\text{log-likelihood ratio}} \tag{6}$$

By letting α_i differ from 1, we allow each subject i to be more or less responsive to information than a Bayesian. We call α_i subject i's *responsiveness to information*.

Our model of responsiveness to information is equivalent to applying a linear-in-log-odds probability weighting function (Goldstein and Einhorn, 1987; Gonzalez and Wu, 1999; Tversky and Fox, 1995) to the Bayesian posterior. Formally, subject i's posterior beliefs and the Bayesian posterior $P_I(s = 1|\sigma)$ are related by[26]

$$P_{I,i}(s = 1|\sigma) = \frac{\delta_i P_I(s = 1|\sigma)^{\alpha_i}}{\delta_i P_I(s = 1|\sigma)^{\alpha_i} + \big(1 - P_I(s = 1|\sigma)\big)^{\alpha_i}} \tag{7}$$

with $\delta_i = \left(\frac{P(1)}{1-P(1)}\right)^{1-\alpha_i}$ and $P(1)$ denoting prior beliefs. This function intersects the 45°-line at the prior $P(1)$, and has the familiar inverse S-shape if $0 < \alpha_i < 1$. (See Fig. 4.)

[23] We control for order and session fixed effects. To perform this exercise with $\Delta P_{I,i}(s = 0|\sigma = 0)$ we drop the boundary information structures, and split the sample in pieces of two and five information structures. This is because $\text{logit}(x)$ is undefined for $x = 1$. We bootstrap the standard errors by drawing 1000 bootstrap samples (block-bootstrapped on individuals).

[24] This is robust to alternative methods of measuring consistency, such as the Cronbach alpha statistic, which is 0.915, 0.912 and 0.870 for $\Delta v_{I,i}$, $\Delta P_{I,i}(s = 1|\sigma = 1)$, and $\Delta P_{I,i}(s = 0|\sigma = 0)$, respectively. These compare favorably with the standard benchmark of 0.8 (Kline, 1999), indicating a high consistency of behavior within parts of the experiment.

[25] These authors did not examine individual consistency.

[26] To see this, write equation (6) both for subject i and for a Bayesian, and then substitute out the log-likelihood ratio.

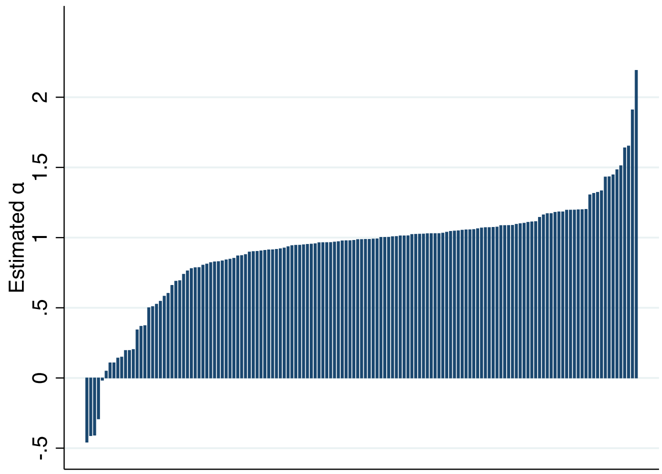


Fig. 5. The estimated responsiveness parameters $\hat{\alpha}_i$ by subject. Subjects ordered by $\hat{\alpha}_i$.

Our model is *not* identical to standard probability weighting (as in Kahneman and Tversky, 1979 or Tversky and Kahneman, 1992). Standard theories of probability weighting do not distinguish between lotteries with explicitly stated distributions and lotteries requiring probabilistic inference. Thus, they predict no divergence from Bayesian behavior in the belief updating task.[27] By contrast, our model involves applying a probability-weighting-style transformation *just to the Bayesian posterior beliefs*, and those divergences from Bayesian behavior are the core of our empirical strategy to predict behavior in other tasks.

We estimate each subject's parameter of responsiveness to information using *only* the data from the Belief Updating Task. We estimate the following regression model using generalized least squares.[28]

$$\begin{bmatrix} \text{logit}(P_{I,i}(s=1|\sigma=1)) \\ \text{logit}(P_{I,i}(s=0|\sigma=0)) \end{bmatrix} = \alpha_i \begin{bmatrix} \log\left(\frac{P_I(\sigma=1|s=1)}{P_I(\sigma=1|s0)}\right) \\ \log\left(\frac{P_I(\sigma=0|s=0)}{P_I(\sigma=0|s=1)}\right) \end{bmatrix} + \begin{bmatrix} \epsilon^1_{I,i} \\ \epsilon^0_{I,i} \end{bmatrix} \tag{8}$$

We do not use data on the boundary information structures because $\text{logit}(P_{I,i}(s=0|\sigma=0))$ is not defined when $P_{I,i}(s=0|\sigma=0)=1$. Logit priors vanish because $P(s=1)=1/2$.[29]

Fig. 5 displays the distribution of the estimated parameters $\hat{\alpha}_i$. Our average subject is slightly conservative – the mean $\hat{\alpha}_i$ is 0.914. The $\hat{\alpha}_i$ are quite variable across subjects, with a standard deviation of 0.396, a minimum of -0.457 and a maximum of 2.191. We reject the null hypothesis that $\hat{\alpha}_i = 1$ at the 5% level for 43% (62 of 143) subjects.

4.2. Explanatory power across environments

Setting 1: Information Valuation Task If responsiveness to information is a *stable* individual trait, then our individual-level estimates of responsiveness should explain behavior in the Information Valuation Task (an environment which we have *not* used to estimate individual responsiveness parameters). If responsiveness to information is an *important* trait in our setting, then we should not gain much explanatory power by using the belief updating data in a way that allows for deviations that are idiosyncratic to information structures, rather than by using responsiveness alone.

Subject behavior is correlated across the Information Valuation Task and the Belief Updating Task. The partial correlation between $\Delta v_{I,i}$ and $\Delta P_{I,i}(s=0|\sigma=0) + \Delta P_{I,i}(s=1|\sigma=1)$ (controlling for session and order fixed effects) is 0.328 ($p < 0.0001$). Hence, a subject's valuations of each information structure are substantially explained by her posterior beliefs for that information structure, *even after we control for informativeness*.

To investigate what part of the variation in information demand responsiveness can explain, we use the data from the Belief Updating Task to construct two distinct predictors of information demand behavior, and compare their explanatory power.

Our *constrained predictor* $\Delta v_I(\hat{\alpha}_i)$ uses *only a subject's estimated parameter of responsiveness* to predict how far her valuations deviate from the Bayesian benchmark. Since responsiveness to information is a single parameter, this prevents

[27] In the Belief Updating task, subjects choose between lotteries that pay \$35 with a stated probability and \$0 otherwise, and lotteries that pay \$35 if a given state was realized and \$0 otherwise. In standard probability weighting, a subject who believes that the given state was realized with probability 70% is indifferent between a lottery that pays \$35 if that state was realized and an exogenous lottery with a 70% chance of \$35.

[28] We use GLS rather than OLS, because $\epsilon^1_{I,i}$ and $\epsilon^0_{I,i}$ are plausibly correlated for any given information structure. Hence GLS is a more efficient estimator.

[29] Responsiveness to information has explanatory power even in the Gradual Information Task, in which subjects generally update from non-uniform priors.

Table 4
Explaining and decomposing deviations from benchmark valuations. Significance levels on the slope parameters concern the hypothesis that the coefficients are 1. Standard errors are bootstrapped using the following semiparametric procedure: For each subject we first draw a responsiveness parameter α_i according to the distribution of his estimate of this parameter. We then select a bootstrap sample and perform the estimations. We repeat this 1000 times. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variables	(1) $\Delta v_{I,i}$	(2) $\Delta v_{I,i}$	(3) $\Delta v_{I,i}$	(4) $\Delta v_{I,i}$	(5) $\Delta v_{I,i}$	(6) $\Delta v_{I,i}$
$\Delta v_{I,i}^{Beliefs}$	0.246^{***}		0.413^{***}	0.214^{***}		
	(0.066)		(0.083)	(0.080)		(0.092)
$\Delta v_I(\hat{\alpha}_i)$	0.300^{***}	0.550^{***}		1.020	1.244	
	(0.146)	(0.144)		(0.359)	(0.360)	
Constant	-0.038^{***}	-0.035^{***}	-0.044^{***}	-0.039^{***}	-0.036^{***}	-0.049^{***}
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Obs.	1,429	1,429	1,429	1,119	1,119	1,119
R^2	0.130	0.113	0.116	0.107	0.093	0.056
10% trim	–	–	–	Yes	Yes	Yes

information-structure-dependent effects, unless they are implied by the responsiveness model. Formally, for each subject and information structure, we use that subject's $\hat{\alpha}_i$ to predict her valuations $\hat{v}_I(\hat{\alpha}_i) = P_I(\sigma = 1)\hat{P}_I(s = 1|\sigma = 1; \hat{\alpha}_i) + P_I(\sigma = 0)\hat{P}_I(s = 0|\sigma = 0; \hat{\alpha}_i)$, where, $\hat{P}_I(s|\sigma; \hat{\alpha}_i)$ are the predicted posteriors. We then define $\Delta v_I(\hat{\alpha}_i) := \hat{v}_I(\hat{\alpha}_i) - v_I^{Bayes}$.

Our *unconstrained predictor* $\Delta v_{I,i}^{Beliefs}$ uses each subject's posterior beliefs to predict how far her valuation of an information structure will deviate from the Bayesian benchmark *separately for each information structure*. This permits idiosyncratic deviations for each information structure. Hence, it may be influenced by many distinct belief updating biases along which individuals may be heterogeneous, such as different insensitivity to signal characteristics, different rules of thumb, as well as functional form misspecification of the responsiveness model. Formally, $\Delta v_{I,i}^{Beliefs} := v_{I,i}^{Beliefs} - v_I^{Bayes}$.[30]

We compare the explanatory power of the two predictors by estimating

$$\Delta v_{I,i} = \beta_0 + \beta_1 \Delta v_{I,i}^{Beliefs} + \beta_2 \Delta v_I(\hat{\alpha}_i) + \epsilon_i \tag{9}$$

We interpret the R^2 of this model as the *maximal* part of the variance in the valuations of information structures that can be explained by belief updating. To assess how well responsiveness to information explains this variance, we estimate model (9) with β_1 restricted to 0 and study the drop in R^2.

The results are in Columns 1 and 2 of Table 4.[31] The R^2 of the unconstrained model is 13.0%, which compares to an R^2 of 11.3% for the constrained model.[32] This shows that our simple one-parameter model captures most (86.9%) of the variation in valuations that can be explained by non-standard belief updating.[33] Once we know a subject's general responsiveness to information, a subject's specific *ex post* beliefs contribute little explanatory power. For completeness, Column 3 reports the estimates of model (9) when β_2 is set to 0.[34] Our results are not driven by outliers. Columns 4–6 show that once we drop the 20% of subjects with the most extreme estimated parameters of responsiveness to information, the one-parameter model still explains 86.9% of the variation in valuations that can be explained by non-standard belief updating.

It bears emphasis that the $\Delta v_I(\hat{\alpha}_i)$ variables are generated using only data from the Belief Updating Task. They obtain explanatory power in the Information Valuation Task even though they are generated without using data from that task.[35]

Setting 2: Gradual Information Task The Gradual Information Task is a second environment that allows us to test the robustness of responsiveness to information. Subjects observed a *sequence* of informative signals about the state of the world, and made choices that revealed when their certainty about the underlying state crossed a given threshold. Hence while the Belief Updating Task reveals how certain a subject is about the state of the world after observing a signal realization, the Gradual Information Task reveals how much information a subject needs to observe before attaining a given level of certainty. Additionally, unlike in the Belief Updating Task, subjects in the Gradual Information Task (i) generally update

[30] Alternatively, one could exclude data from the Signal Probability Assessment task, and construct $v_{I,i}^{Beliefs2} = P(\sigma = 1)P_{I,i}(s = 1|\sigma = 1) + P(\sigma = 0)P_{I,i}(s = 0|\sigma = 0)$. This specification uses the *true* signal probabilities and the elicited posterior beliefs. The results that follow are essentially unchanged if we use this specification instead.

[31] Since we use estimates of α_i as explanatory variables, we bootstrap the standard errors. We use the following semiparametric procedure: For each subject we first draw a responsiveness parameter α_i according to the distribution of his estimate of this parameter. We then select a bootstrap sample and perform the estimations. We draw a total of 1000 bootstrap samples.

[32] The difference between these R^2 values is significant at all conventional levels. (Testing that the R^2 coefficient of the model in column 1 is significantly larger than in the model in column 2 is equivalent to testing that the coefficient on $\Delta v_{I,i}^{Beliefs}$ is zero in the model in column 1, see Greene, 2008, chapter 5.)

[33] We show in Appendix B.4 that this is not driven by a particular set of information structures.

[34] Notably, the R-squared in this instance is lower than that of the *constrained* model. This suggests that the constrained model exploits underlying structure in the data – the stability of behavior across choice problems in the same task – to produce better estimates.

[35] Moreover, the R^2 coefficients in Table 4 are deflated by noise in the data. If we average $\Delta v_{I,i}$, $\Delta v_{I,i}^{Beliefs}$, and $\Delta v_I(\alpha_i)$ across information structures, we obtain R^2 coefficients of 21.6, 16.7 and 20.0% in the regressions corresponding to columns 1–3 of Table 4, respectively.

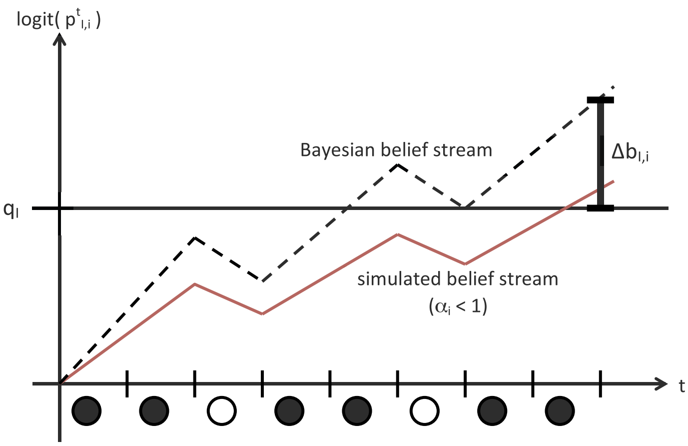


Fig. 6. Evolution of Bayesian beliefs and those of a subject with responsiveness $\alpha_i < 1$. The subject requires more information than the Bayesian before her beliefs cross q_i.

Table 5
Predicting behavior in the Gradual Information Task. Results are from Tobit-regressions. Demographic controls include race and gender. For goodness-of-fit, we report the McKelvey and Zavoina (1975) pseudo-R^2. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variables	(1) $\Delta b_{I,i}$	(2) $\Delta b_{I,i}$	(3) $\Delta b_{I,i}$	(4) $\Delta b_{I,i}$
$\hat{\alpha}_i$	-1.65^{**}	-1.77^{**}		
	(0.79)	(0.75)		
$\Delta b_I(\hat{\alpha}_i)$			0.23	0.22
			(0.19)	(0.18)
Constant	4.03	3.80^{***}	2.28^{***}	2.08^{***}
	(0.86)	(1.00)	(0.38)	(0.68)
Demographic controls	No	Yes	No	Yes
Observations	572	572	572	572
Pseudo-R^2	0.028	0.032	0.014	0.013

from non-symmetric priors (after having observed the first piece of information), (ii) make aggregative assessments from a long sequence of signals, (iii) make ternary decisions rather than an essentially continuous choice, (iv) choose after having observed information rather than conditional on observing information in the future.

Consider how a subject perfectly described by the responsiveness model would behave in the Gradual Information Task. Her posterior belief $p^t_{I,i}$ that $s = 1$ after having observed t conditionally independent signal realizations is derived by repeated application of equation (6). Thus, her logit-beliefs thus follow a random walk with drift, where the responsiveness parameter α_i scales the rate of movement. Hence, we expect individuals who are *less* responsive to information to require *more* ball draws before they are willing to bet on a box, and vice versa. Fig. 6 illustrates.

Our variable of interest is how much more (or less) sure than q_I a Bayesian would be at the point in time $\tau_{I,i}$ at which subject i first bets on a box. We define $\Delta b_{I,i} = \text{logit}(b^{\tau_{I,i}}_{I,i}) - \text{logit}(q_I)$ where $b^{\tau_{I,i}}_{I,i}$ is the Bayesian posterior at $\tau_{I,i}$ for the sequence of information that subject i has observed.[36] This is marked on Fig. 6. The responsiveness model predicts that $\Delta b_{I,i}$ is negatively related to α_i.

We regress $\Delta b_{I,i}$ on each subject's responsiveness $\hat{\alpha}_i$, which is estimated using only data from the Belief Updating Task. Columns 1 and 2 in Table 5 report the results.[37] Responsiveness has the expected effect on subject performance in the Gradual Information Task. This effect is significant with and without the inclusion of demographic controls.

To take a fully structural approach, we take each subject's responsiveness parameter, and generate a prediction for Δb_I given the sequence of signals each subject saw. We call this variable $\Delta b_I(\hat{\alpha}_i)$. We then regress measured $\Delta b_{I,i}$ on the prediction $\Delta b_I(\hat{\alpha}_i)$. Columns 3 and 4 in Table 5 report the results. The coefficient has the expected sign, but is not

[36] Alternatively, one could consider the number of balls that a subject saw before betting on a box. Since ball draws are conditionally i.i.d., however, this variable is noisy. Applying the logit transformation to beliefs makes our measure comparable across different information structures and ensures that overly and underly responsive subjects receive the same weight in the regression.

[37] Our data are truncated, since the random drawing of balls caused some subjects to not observe a sufficient amount of information that would have made them willing to bet on the state of the world. Consequently, we use Tobit regressions.

Table 6
Cognitive style, mathematical aptitude and responsiveness to information. The excluded category for college major includes all non-STEM fields that are neither business nor economics. Columns 1 and 3 list the results of OLS regressions, columns 2 and 4 those of median regressions. All models estimated with session and order fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variables	(1) α_i	(2) α_i	(3) $\lvert\alpha_i - 1\rvert$	(4) $\lvert\alpha_i - 1\rvert$
Knows Bayes' law	0.098	−0.021	0.013	0.003
	(0.081)	(0.053)	(0.065)	(0.026)
CRT score	0.051	0.047*	−0.039	−0.027**
	(0.038)	(0.025)	(0.031)	(0.012)
STEM major	−0.044	−0.008	−0.053	−0.021
	(0.078)	(0.050)	(0.063)	(0.025)
Business or economics major	−0.142	−0.075	0.016	0.005
	(0.133)	(0.084)	(0.107)	(0.043)
Demographic controls	Yes	Yes	Yes	Yes
Method	OLS	QR	OLS	QR
Constant	0.892***	0.785***	0.449***	0.294***
	(0.131)	(0.081)	(0.105)	(0.041)
Observations	143	143	143	143
R-squared	0.143	–	0.155	–

statistically significant. This suggests that, while measured responsiveness predicts behavior in different choice problems, the functional form we assumed in Equation (6) may not straightforwardly carry over to the Gradual Information Task.

Alternative explanations We have elicited subjects' posterior beliefs and valuations of information structures using multiple price lists. These are a specific implementation of the Becker et al. (1964) mechanism, which experimental subjects sometimes find hard to understand (e.g. Andersen et al., 2006 and Beauchamp et al., 2015).

If there is consistent individual heterogeneity in subjects' biases in this elicitation procedure, those potentially confound the effects of responsiveness to information. We address this concern in two ways. First, if our results were an artifact of consistent individual heterogeneity in misunderstandings of multiple price list procedures, we would not expect to find that our estimates of responsiveness are correlated with behavior in the gradual information task, which does not involve multiple price lists.

Second, regarding consistency between the belief updating and information valuation tasks, our design lets us difference out the effect of consistent individual heterogeneity in the misunderstanding of multiple price lists. The reason is that misunderstanding of multiple price lists may affect behavior in all three tasks that rely on multiple price lists, specifically the information valuation, belief updating, and probability assessment tasks. Responsiveness to information, by contrast, may affect behavior only if belief updating plays a role; specifically in the belief updating and information valuation tasks. Statistically significant and substantial correlations remain, showing that individual consistency does not (only) derive from misunderstandings of the multiple price lists. We defer the formal analysis to Appendix B.3.[38]

4.3. Correlation with other variables

In a cognitively demanding experiment such as ours, behavior may be caused by more general stable individual traits such as mathematical aptitude and cognitive style. While we do not have an *ex ante* hypothesis whether more mathematically proficient subjects would be more or less responsive to information, one might expect that they would differ from the Bayesian benchmark to a smaller extent. Our data do not suggest any such correlation.

Formally, we regress both the estimated parameters of responsiveness and their deviation from the Bayesian benchmark on subjects' self-reported knowledge of Bayes' law, their cognitive style as measured by the CRT score, their college major, and a vector of demographic characteristics consisting of dummies for race and gender.

Column 1 of Table 6 shows that neither of our measures of mathematical aptitude and cognitive style is significantly associated with responsiveness to information. Once we use median regression instead of least squares (Column 2), we find a significantly positive relation between CRT scores and the subjects' responsiveness. We reach similar conclusions by associating these variables with the extent to which a subject deviates from the Bayesian benchmark. Using OLS, we find no

[38] Another potential explanation for individual consistency concerns the different random draws of balls that different subjects were exposed to in part 1 of the experiment. These varying experiences could simultaneously affect all choices relating to random ball draws, including the probability assessment task. Accordingly, we can difference out such effects. The consistency in behavior across the belief updating and information valuation tasks remains. See Appendix B.3 for details.

Table 7
Relating the responsiveness model to the information-structure-specific effects. Reported significance levels on v_I^{Bayes} concern the hypothesis that the coefficient is 1. Standard errors clustered by subject. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variables Sample	(1) $v_I(\alpha_i)$ all	(2) $v_{I,i}$ all	(3) $v_I(\alpha_i)$ all	(4) $v_{I,i}$ all	(5) $v_I(\alpha_i)$ $\alpha_i < 1$	(6) $v_{I,i}$ $\alpha_i < 1$	(7) $v_I(\alpha_i)$ $\alpha_i > 1$	(8) $v_{I,i}$ $\alpha_i > 1$
v_I^{Bayes}	0.883***	0.603***	0.877***	0.615***	0.751***	0.560***	1.020***	0.677***
	(0.026)	(0.043)	(0.027)	(0.041)	(0.046)	(0.060)	(0.003)	(0.054)
a_I	−0.007**	−0.024	0.023***	−0.090	0.030***	−0.060	0.013***	−0.126
	(0.003)	(0.026)	(0.005)	(0.075)	(0.008)	(0.105)	(0.002)	(0.106)
$(a_I)^2$			−0.084***	0.190	−0.177***	−0.069	0.023***	0.485
			(0.020)	(0.198)	(0.033)	(0.248)	(0.002)	(0.315)
b_I	0.022***	0.034***	0.023***	0.032***	0.056***	0.060***	−0.015***	0.002
	(0.006)	(0.010)	(0.006)	(0.010)	(0.009)	(0.015)	(0.002)	(0.012)
Constant	0.062***	0.260***	0.065***	0.252***	0.115***	0.253***	0.010**	0.251***
	(0.013)	(0.033)	(0.013)	(0.031)	(0.023)	(0.044)	(0.004)	(0.044)
Observations	1,429	1,429	1,429	1,429	760	760	669	669
R^2	0.507	0.177	0.508	0.177	0.436	0.179	0.925	0.220

significant relations (Column 3), but once we estimate median regressions, we discover that subjects who score one point higher in the CRT test are in closer alignment with the Bayesian benchmark by 2.6 percentage points (Column 4).[39]

4.4. Responsiveness and information-structure-specific biases

We now tie our one-parameter model to the information-structure-specific biases that we documented in section 3. Such a relation is perhaps surprising – after all, the responsiveness model is set up to capture individual-level biases that do not depend on the specifics of information structures.

The model predicts that under-responsive subjects will exhibit both the compression effect and the certainty effect. An under-responsive subject still correctly identifies wholly uninformative signals. However, as a signal becomes marginally more informative, an under-responsive subject's beliefs move *less* than a Bayesian's would, which means that his valuations also increase less than Bayesian's. This generates the compression effect. The certainty effect arises because a subject also has correct beliefs upon observing a perfectly informative signal realization. Therefore, for a boundary information structure, the individual will have too low a posterior only for the signal realization that leaves residual uncertainty, but not for the one that reveals the state of the world. For an interior information structure, by contrast, the individual will have too low a posterior regardless of the signal realization.

Formally, we derive these predictions from the probability-weighting interpretation of the responsiveness model when $0 < \alpha_i < 1$. The compression effect is due to the fact that the slope of (7) is lower than 1 for all Bayesian posteriors that are not too close to 0 or 1. Moreover, this function is convex to the right of the prior, and most strongly so for Bayesian posteriors close to 1. Hence it implies the certainty effect (as well as a more general but weak preference for asymmetric information structures).

Since the majority of our subjects in our sample are less responsive than a Bayesian, we expect that the responsiveness model with the estimated parameters will reproduce the information-structure-specific biases. For each subject i and information structure I we use the estimated parameter α_i to predict the distance between that subjects' valuation of that information structure and the Bayesian benchmark. We then estimate model (4) with the predicted deviations as the dependent variable. Table 7, column 1 reports the results. For comparison, column 2 reproduces the analysis with the measured deviations as the dependent variable. The coefficient on v_I^{Bayes} is significantly smaller than 1, and hence the model captures the compression effect. The extent of this effect, however, is significantly smaller than in column 2. The predicted valuations also exhibit a significant certainty effect. Again, the extent of the bias in the predicted data is smaller than the measured one.

The coefficient on asymmetry is negative even for the predicted valuations, in contrast to the theoretical predictions. This indicates that the linear functional form may be overly restrictive. Thus we include squared asymmetry as a regressor and report the results in columns 3 and 4. Estimates of the compression and certainty effects do not substantially change for measured, nor predicted, valuations.

The mechanics that link the responsiveness model to the information-structure-specific biases rest on the fact that the average subject is less responsive than a Bayesian. For overly responsive subjects, the information-structure-specific biases are predicted to reverse in sign. To test this hypothesis, we reproduce the above analysis separately on the subsamples of subjects with $\alpha_i < 1$ and those with $\alpha_i > 1$. Columns 5–8 display the results. (The dependent variable is the predicted deviation in columns 5 and 7, and the measured deviation in columns 6 and 8.) While we do not see the predicted reversal of the compression and certainty effects for overly responsive subjects, we do find that both the compression effect and the

[39] In line with these findings, Stanovich and West (2008) conclude that for the type of belief updating problems we study, discrepancies between normative models and measured behavior are not easily explained by recourse to cognitive capacity.

certainty effect are significantly stronger for the subjects with $\alpha_i < 1$ than for the remaining subjects ($p < 0.01$ for each of the effects).

5. Conclusion

Economists study many situations where agents learn about their environment by acquiring costly stochastic information. In this article, we examine information demand behavior and its relation to belief updating and non-standard risk preferences across a range of choice contexts.

We find two biases that depend on the properties of information structures. First, subjects generally underreact to increases in the informativeness of an informative signal. Second, they value boundary information structures disproportionately highly. These two biases do not persist, however, once we condition on subjects' elicited beliefs, rather than the objective Bayesian posteriors. Hence, these biases are primarily the result of non-standard belief updating. They are unlikely caused by non-standard risk preferences.

Importantly, we identify a bias that is specific to individuals. They differ *consistently* in responsiveness to information, the extent to which their beliefs move upon observing information. Individual-level estimates of responsiveness to information have explanatory power in two different choice problems. They explain about 80% of the variation in information demand that is maximally attributable to belief updating. Responsiveness is unrelated to proxies for mathematical aptitude. The one-parameter model of responsiveness to information we estimate is also able to generate the compression and certainty effects, biases that depend on the properties of information structures. Intuitively, this arises because our average subject is less responsive than a Bayesian, and such subjects' beliefs move too little for signals of intermediate informativeness, but are in line with the benchmark for perfectly informative and for perfectly uninformative signals.

If we use our individual estimates of responsiveness in a structural model, however, they fail to predict behavior in our sequential belief updating task with statistical significance (although the raw coefficients are significantly correlated with behavior in that task, and even the estimates from the structural model attain the predicted sign). This opens interesting questions for further research. Specifically, the raw parameters are more correlated with behavior in that task than the structural predictions that rely on them. This suggests that the structural model is *correctly identifying* a dimension of individual heterogeneity, but *incorrectly predicting* how this heterogeneity manifests in different tasks. In particular, if a single signal and a sequence of signals contain the same information, the responsiveness model constrains agents to evaluate both identically. However, it could instead be that individuals over-infer from short sequences of signals, but under-infer from long ones (Griffin and Tversky, 1992; Kraemer and Weber, 2004), so a model that predicts well across different kinds of belief-updating tasks may have to abandon this invariance.

More generally, the finding highlights the importance of determining the scope of applicability of models of belief updating. Two seemingly opposite conclusions emerge from prior work on belief updating. An early psychology literature employed balls-and-urn tasks, and documented conservatism, *i.e.* that subjects update too little from signals (see Peterson and Beach (1967) for a review). The heuristics and biases approach, by contrast, has often relied on tasks that do not provide enough formal structure to allow for the explicit calculation of Bayesian posterior beliefs, and has often documented base-rate neglect, which leads subjects to update too much from weak signals (for instance the Tom W. example in Kahneman and Tversky, 1972). In our data, we find that subjects value signals of varying informativeness as though they were more alike, compared to the predictions of the standard model (the compression effect). An expanded model that exploits this observation might accommodate both conservatism and base-rate neglect.

Moreover, the tasks in our experiment are all based on a similar balls-and-urns setting. It would be valuable for future research to explore how predictive responsiveness to information is across more radically different environments. For instance, are subjects' choices in our experimental tasks related to the extent to which they update upon receiving the results of a medical test, or of the extent to which they change their financial investments upon learning financial news? Recent research suggests that such an investigation holds promise. On the one hand, Buser et al. (2016) find that responsiveness to information predicts tournament entry. On the other hand, that the geopolitical forecasters studied in Augenblick and Rabin (2015) exhibit individual consistency in the volatility of belief streams.[40]

Responsiveness to information also bears potential for theoretical generalization. Our application studies the two-states case, in which posterior variance is a deterministic function of the posterior mean. In the more general case, does responsiveness operate on the dimension of means, variance, or both? That is, in the language of Moore and Healy (2008), is an overly responsive individual prone to overplacement, to overprecision, or to both?

Once individual stability and predictiveness is further established, and more responsive individuals can be identified, economists may begin using the trait to better target programs such as those fostering technology adoption in developing countries. Relatedly, in settings where non-Bayesian updating is costly, such as medical decision making, subjects who are excessively or insufficiently responsive to information might be identified and given help to bring their decision-making in line with the normative benchmark.

[40] To see the connection between responsiveness and volatility, consider two agents who see the same stream of information about a binary event. Suppose one of them is extremely responsive whereas the other's responsiveness is close to zero. The former agent's beliefs will tend to jump back and forth between values close to 0 and close to 1, whereas the latter agent's beliefs will remain close to his prior.

Acknowledgments

We are deeply indebted to our advisors B. Douglas Bernheim, Paul Milgrom, Muriel Niederle, Alvin E. Roth, and Charles Sprenger. We are grateful to Itay Fainmesser, Simon Gaechter, Mark Machina, Jeffrey Naecker, Joel Sobel and participants at the Stanford Behavioral and Experimental Economics Lunch, the conference of the Economic Science Association, the conference of Swiss Economists Abroad, the Bay Area Behavioral and Experimental Economics Workshop, and the Stanford Institute for Theoretical Economics for very helpful comments and discussions.

Funding: This work was supported by the Department of Economics at Stanford University. This research was approved by the Stanford IRB in protocol 27326.

Appendices A–D. Supplementary material

Supplementary material Appendices A–D related to this article can be found online at https://doi.org/10.1016/j.geb.2017.11.009.

References

Abrevaya, J., Jiang, W., 2005. A nonparametric approach to measuring and testing curvature. J. Bus. Econ. Statist. 23 (1), 1–19.
Allen, F., 1987. Discovering personal probabilities when utility functions are unknown. Manage. Sci. 33 (4), 307–319.
Andersen, S., Harrison, G.W., Lau, M.I., Rutstrom, E.E., 2006. Elicitation using multiple price list formats. Exper. Econ. 9, 383–405.
Antoniou, C., Harrison, G.W., Lau, M., Read, D., 2017. Information characteristics and errors in expectations: experimental evidence. J. Finan. Quant. Anal. 52 (2), 737–750.
Athey, S., Levin, J., 2001. The value of information in monotone decision problems. Stanford University. Mimeo.
Augenblick, N., Rabin, M., 2015. Testing for excess movement in beliefs. Mimeo.
Azrieli, Y., Lehrer, E., 2008. The value of a stochastic information structure. Games Econ. Behav. 63, 679–693.
Baillon, A., Bleichrodt, H., 2015. Testing ambiguity models through the measurement of probabilities for gains and losses. Am. Econ. J. Microecon. 7 (2), 77–100.
Bastardi, A., Shafir, E., 1998. On the pursuit and misuse of useless information. J. Pers. Soc. Psychol. 75 (1), 19–32.
Beauchamp, J.P., Benjamin, D.J., Chabris, C.F., Laibson, D.I., 2015. Controlling for the Compromise Effect Debiases Estimates of Risk Preference Parameters. Tech. rep. National Bureau of Economic Research.
Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behav. Sci. 9 (3), 226–232.
Benjamin, D.J., Rabin, M., Raymond, C., 2015. A model of nonbelief in the law of large numbers. J. Eur. Econ. Assoc. 14 (2), 518–544.
Bernasconi, M., Loomes, G., 1992. Failures of the reduction principle in an Ellsberg-type problem. Theory Dec. 32 (1), 77–100.
Brown, M., Flinn, C.J., Schotter, A., 2011. Real-time search in the laboratory and the market. Amer. Econ. Rev. 101 (2).
Burks, S.V., Carpenter, J.P., Goette, L., Rustichini, A., 2013. Overconfidence and social signaling. Rev. Econ. Stud. 80 (3), 949–983.
Buser, T., Gerhards, L., van der Weele, J., 2016. Measuring Responsiveness to Feedback as a Personal Trait. Working paper.
Cabrales, A., Gossner, O., Serrano, R., 2013a. The Appeal of Information Transactions. Working paper.
Cabrales, A., Gossner, O., Serrano, R., 2013b. Entropy and the value of information for investors. Amer. Econ. Rev. 103 (1), 360–377.
Camerer, C., 1995. Individual decision making. In: Kagel, J.H., Roth, A.E. (Eds.), Handbook of Experimental Economics. Princeton University Press, pp. 587–683. Ch. 8.
Caplin, A., Dean, M., Martin, D., 2011. Search and satisficing. Amer. Econ. Rev. 101 (7).
Copeland, T., Friedman, D., 1987. The effect of sequential information arrival on asset prices: an experimental study. J. Finance 42.
Eil, D., Rao, J.M., 2011. The good news–bad news effect: asymmetric processing of objective information about yourself. Am. Econ. J. Microecon. 3, 114–138.
El-Gamal, M.A., Grether, D.M., 1995. Are people Bayesian? Uncovering behavioral strategies. J. Amer. Statistical Assoc. 90, 432.
Eliaz, K., Schotter, A., 2010. Paying for confidence: an experimental study of the demand for non-instrumental information. Games Econ. Behav. 70, 304–324.
Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. Quart. J. Econ., 643–669.
Erev, I., Haruvy, E., 2015. Learning and the economics of small decisions. In: Kagel, J.H., Roth, A.E. (Eds.), The Handbook of Experimental Economics, vol. 2.
Ergin, H., Gul, F., 2009. A theory of subjective compound lotteries. J. Econ. Theory 144, 899–929.
Falk, A., Zimmermann, F., 2016. Beliefs and utility: experimental evidence on preferences for information. University of Zurich. Mimeo.
Fehr-Duda, H., Epper, T., 2012. Probability and risk: foundations and economic implications of probability-dependent risk preferences. Ann. Rev. Econ. 4, 567–593.
Fiedler, K., Juslin, P. (Eds.), 2006. Information Sampling and Adaptive Cognition. Cambridge University Press.
Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Exper. Econ. 10 (2), 171–178.
Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect., 25–42.
Gabaix, X., Laibson, D., Moloche, G., Weinberg, S., 2006. Costly information acquisition: experimental analysis of a boundedly rational model. Amer. Econ. Rev. 96 (4).
Gilboa, I., Marinacci, M., 2013. Ambiguity and the Bayesian paradigm. In: Advances in Economics and Econometrics: Theory and Applications. Tenth World Congress of the Econometric Society.
Goldstein, W., Einhorn, H., 1987. Expression theory and the preference reversal phenomena. Psychol. Rev. 94, 236–254.
Gonzalez, R., Wu, G., 1999. On the shape of the probability weighting function. Cogn. Psychol. 38, 129–166.
Greene, W.H., 2008. Econometric Analysis, 6th edition. Pearson Prentice Hall.
Grether, D.M., 1980. Bayes rule as a descriptive model: the representativeness heuristic. Quart. J. Econ. 95 (3), 537–557.
Grether, D.M., 1992. Testing Bayes rule and the representativeness heuristic: some experimental evidence. J. Econ. Behav. Organ. 17 (1), 31–57.
Gretschko, V., Rajko, A., 2015. Excess information acquisition in auctions. Exper. Econ. 18 (3), 335–355.
Griffin, D., Tversky, A., 1992. The weighing of evidence and the determinants of confidence. Cogn. Psychol. 24, 411–435.
Halevy, Y., 2007. Ellsberg revisited: an experimental study. Econometrica 75 (2), 503–536.
Halevy, Y., Feltkamp, V., 2005. A Bayesian approach to uncertainty aversion. Rev. Econ. Stud. 72, 449–466.
Hoelzl, E., Rustichini, A., 2005. Overconfident: do you put your money on it? Econ. J. 115 (04), 305–318.
Hoffman, M., 2016. How is information valued? Evidence from framed field experiments. Econ. J. 126, 1884–1911.
Holt, C.A., Smith, A.M., 2009. An update on Bayesian updating. J. Econ. Behav. Organ. 69, 125–134.
Holt, C.A., Smith, A.M., 2016. Belief elicitation with a synchronized lottery choice menu that is invariant to risk attitudes. Am. Econ. J. Microecon. 8 (1).

Kahneman, D., Tversky, A., 1972. Subjective probability: a judgement of representativeness. Cogn. Psychol. 3, 430–454.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 2, 263–292.
Karni, E., 2009. A mechanism for eliciting probabilities. Econometrica 77 (2), 603–606.
Kline, P., 1999. Handbook of Psychological Testing, 2nd edition. Routledge, London and New York.
Knight, F.H., 1921. Risk, Uncertainty, and Profit. Houghton Mifflin Company, Boston, MA.
Kocher, M.G., Krawczyk, M., van Winden, F., 2014. 'Let me dream on!' anticipatory emotions and preference for timing in lotteries. J. Econ. Behav. Organ. 98, 29–40.
Kraemer, C., Weber, M., 2004. How do people take into account weight, strength and quality of segregated vs. aggregated data? Experimental evidence. J. Risk Uncertainty 92 (2), 113–142.
Kubler, D., Weizsacker, G., 2004. Limited depth of reasoning and failure of cascade formation in the laboratory. Rev. Econ. Stud. 71.
Masatlioglu, Y., Orhun, Y., Raymond, C., 2015. Skewness and intrinsic preferences for information. Mimeo.
Massey, C., Wu, G., 2005. Detecting regime shifts: the causes of under- and overreaction. Manage. Sci. 51 (6), 932–947.
McKelvey, R.D., Zavoina, W., 1975. A statistical model for the analysis of ordinal level dependent variables. J. Math. Sociol. 4 (1), 103–120.
Moebius, M.M., Niederle, M., Niehaus, P., Rosenblat, T.S., 2013. Managing self-confidence: theory and experimental evidence. Mimeo.
Moore, D.A., Healy, P.J., 2008. The trouble with overconfidence. Psychol. Rev. 115 (2), 502.
Peterson, C.R., Beach, L.R., 1967. Man as an intuitive statistician. Psychol. Bull. 68 (1), 29–46.
Peterson, C.R., Schneider, R.J., Miller, A.J., 1965. Sample size and the revision of subjective probabilities. J. Exp. Psychol. 69 (5), 522–527.
Roth, A.E., Malouf, M.W., 1979. Game-theoretic models and the role of information in bargaining. Psychol. Rev. 86 (6), 574.
Savage, L.J., 1954. Foundations of Statistics. Wiley, New York.
Schlag, K., van der Weele, J., 2013. Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality. Theor. Econ. Lett. 3 (1), 38–42.
Segal, U., 1987. The Ellsberg paradox and risk aversion: an anticipated utility approach. Int. Econ. Rev. 28, 175–202.
Seo, K., 2009. Ambiguity and second order belief. Econometrica 77 (5), 1575–1605.
Stanovich, K.E., West, R.F., 2008. On the relative independence of thinking biases and cognitive ability. J. Pers. Soc. Psychol. 94 (4), 672–695.
Sunder, S., 1992. Market for information: experimental evidence. Econometrica 60 (3), 667–695.
Szkup, M., Trevino, I., 2015. Costly information acquisition in a speculative attack: theory and experiments. Mimeo.
Tversky, A., Fox, C.R., 1995. Weighting risk and uncertainty. Psychol. Rev. 102, 269–283.
Tversky, A., Kahneman, D., 1992. Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertainty 5 (4), 297–323.
Yates, F., Zukowski, L., 1976. Characterization of ambiguity in decision making. Behav. Sci. 21, 19–25.
Zimmermann, F., 2014. Clumped or piecewise? Evidence on preferences for information. Manage. Sci. 61 (4), 740–753.