Good News and Bad News are Still News: Experimental Evidence on Belief Updating

Alexander Coutts[*][†]

Nova School of Business and Economics

July 17 2017

Abstract

Bayesian updating remains the benchmark for dynamic modeling under uncertainty within economics. Recent theory and evidence suggest individuals may process information asymmetrically when it relates to personal characteristics or future life outcomes, with good news receiving more weight than bad news. I examine information processing across a broad set of contexts: 1) ego relevant, 2) financially relevant, and 3) non value relevant. In the first two cases, information about outcomes is valenced, containing either good or bad news. In the third case, information is value neutral. In contrast to a number of previous studies I do not find differences in updating patterns across valenced and value neutral settings. While updating across all context and stake conditions is asymmetric and conservative, posteriors remain well approximated by those calculated using Bayes' rule. I find further evidence that asymmetry is sensitive to signal types. Most importantly these patterns are present across all contexts, cautioning against the interpretation of asymmetric updating or other deviations from Bayes' rule as being motivated by psychological biases.

JEL classification: C91, D83, D84.

Keywords: beliefs · Bayes' rule · asymmetric belief updating · conservatism · overconfidence

[*]Nova School of Business and Economics, Faculdade de Economia da Universidade Nova de Lisboa, Campus de Campolide, 1099-032 Lisbon, Portugal; alexander.coutts@novasbe.edu

[†]Acknowledgements: This research has been generously supported by a grant from the Russell Sage Foundation. I am heavily indebted to my advisor David Cesarini for numerous discussions and comments. I am grateful for helpful comments from Hunt Allcott, Kai Barron, Thomas Buser, Colin Camerer, Andrew Demers, David Eil, Guillaume Fréchette, Nicole Hildebrandt, Elliot Lipnowski, David Low, Amnon Maltz, Markus Möbius, Joseph Mullins, Giorgia Romagnoli, Tanya Rosenblat, Andrew Schotter, Emilia Soldani, Tobias Salz, Séverine Toussaert, Joël van der Weele, Christopher Woolnough, Sevgi Yuksel, as well as seminar participants at NYU, the 2015 European ESA Meeeting, the 2016 ASFEE meeting, Games 2016, and THEEM 2017. All errors are my own.

1 Introduction

The ability to process new information in forming and updating beliefs is critical for a wide range of important life decisions. Students receive grades and must adjust their beliefs about their ability to succeed in different subjects before declaring a major, entrepreneurs may be awarded or denied funding for their projects and must update beliefs about the viability of these projects, smokers who are informed of new health statistics on the dangers of smoking must update beliefs about this risk in deciding whether to quit.

In modeling such situations, it is typically assumed that individuals use Bayes' rule to update their beliefs. Individuals who receive partially informative signals about states of the world are assumed to incorporate this information in an unbiased, calculated way. While Bayesian updating is the current paradigm theoretically, there is also a consensus that it has a strong normative basis.

Given the importance of updating beliefs for decision making in economic contexts, experimental evidence on updating has been studied for some time (e.g., Kahneman and Tversky (1973); Grether (1980); Grether (1992); Camerer (1987); Camerer (1995); Holt and Smith (2009)). These studies greatly contributed to our understanding of how individuals update their beliefs, and found important deviations from Bayes' rule. Even so, the nature of the updating tasks in these studies differed considerably from the real updating decisions that motivated them. Unlike updating decisions that are economically relevant to individuals, updating in lab experiments typically involved events such as drawing balls from urns, where subjects hold no personal or financial stake in the outcome, beyond an incentive payment for accuracy. Henceforth, I refer to such probabilistic events as value neutral.

These differences may be critically important, as there is now a small but growing body of theory and empirical evidence suggesting that information is processed differently in non-value neutral, i.e. valenced contexts, depending on whether it is perceived as good or bad news (e.g., Eil and Rao (2011); Sharot et al. (2011); Ertac (2011); Mobius et al. (2014); Kuhnen (2014)). Drawing in part on this evidence, Sharot et al. (2012) claim: "Humans form beliefs asymmetrically; we tend to discount bad news but embrace good news."

In this paper I conduct an experiment to directly investigate whether asymmetric updating is a robust phenomenon of human information processing, by studying the mechanics of belief updating across different contexts and stake conditions. Moving beyond single context studies, I examine binary events that are either 1) ego relevant, 2) financially relevant, or 3) non value relevant. Comparing non value relevant to value contexts generates a robust set of counterfactual updating decisions, overcoming potentially problematic comparisons of previous studies. Further I utilize sizeable financial stakes in order to ensure salience, as a substantial proportion of subjects have the chance to earn $80 in the experiment.

To fix terms, value relevant events are those in which an individual strictly prefers one outcome to another, and news about the outcome can be categorized as good or bad. Non value relevant events are those in which individuals are indifferent between the two outcomes, hence information is just news, neither good nor bad. As this experiment utilizes a financially incentivized belief elicitation procedure, this introduces a type of financial stake. To minimize confusion, I refer to this type of financial stake as an accuracy payment.

In the experiment, individuals receive partially informative binary signals regarding the outcomes of these events. I elicit beliefs utilizing the incentive compatible elicitation procedure of Grether (1992), Holt and Smith (2009), and Karni (2009), and follow Mobius et al. (2014) in estimating an empirical model of belief updating that nests Bayesian updating as a special case, but allows for differential response to affirmative versus negative signals. The elicitation procedure improves on previous work that utilizes other elicitation procedures, such as the quadratic scoring rule (QSR), that are not invariant to subjects' risk preferences.[1]

The results show that, common to previous studies, updating behavior deviates from the strict Bayesian model. In value relevant contexts, updating is conservative, with many non-updates, and asymmetric, with negative signals receiving more weight than affirmative signals. Yet critically, these same deviations are observed in non value relevant contexts, i.e. regardless of whether signals contain good or bad news, or are simply conveying neutral information. Finally, despite these deviations, posteriors are well approximated by those calculated using Bayes' rule.

Overall the analysis indicates the importance of observing a broad set of counterfactual belief updates, as further results of this paper show that pairwise event-comparisons of updating may result in misleading conclusions. Moreover, it suggests caution in attributing biased updating patterns to contexts where such bias is psychologically plausible, as updating patterns are similar across settings where such bias is clearly implausible. The remainder of the paper is as follows. The following section discusses recent theoretical and empirical work investigating belief updating. Next, I outline the experimental design, followed by a description of the results, concluding with a brief discussion.

[1] Antoniou et al. (2015) discuss issues that may arise with inference of updating behaviors when elicitation procedures are not robust to risk preferences.

2 Related Literature

This paper is related to a sizeable literature that exists on studying how individuals process information and whether this is well approximated by Bayes' rule. The majority of these studies investigate updating about non-value relevant events, where subjects have no personal or financial stake in the outcome, excepting a payment for accurate belief reports. Early studies of updating found biases in updating behavior which involved putting too much or too little weight on the prior, and some evidence for a representativeness bias.[2]

Overall, the evidence suggests that individuals do not appear to follow the exact mechanics of Bayes' rule. Yet across contexts, Bayesian approximation does relatively well, and fewer deviations are observed for more experienced subjects, see Camerer (1987) and Camerer (1995). More recently, Holt and Smith (2009) find that individual updating is consistent with Bayes' rule, however they find systematic deviations for extreme values of the prior as well as some evidence that individuals suffer from representativeness bias, consistent with earlier evidence.

More recently, neuroscientists, psychologists, and economists, have argued that individuals might update differently when information is not simply news, but is valenced, in the sense that it may contain good or bad news. There are a number of psychologically plausible motivations for why updating may differ when the context is financially or personally relevant to the individual. In particular, a number of theoretical papers have suggested mechanisms for individuals to update asymmetrically: over-weighting good news relative to bad news. Such information processing may enable individuals to nurture biased beliefs about their abilities or about future outcomes. For example, Landier (2000), Yariv (2005), and Mobius et al. (2014) present models where individuals gain utility from holding positive beliefs, and process information in a biased manner (over-weighting good news relative to bad news) in order to nurture such positive beliefs. Additionally, such biased information processing is similarly rational if optimistic beliefs improve health outcomes, e.g. Scheier and Carver (1987). Finally, biased beliefs could also be nurtured for strategic purposes, as in Benabou and Tirole (2002) regarding self-confidence.

Recognizing this, recent studies have attempted to examine how individuals process information when they have a financial or personal stake in an outcome. In an unincentivized study, Sharot et al. (2011) examined how individuals updated their beliefs about future life

[2]Early work was conducted by Kahneman and Tversky (1973) and Grether (1980), finding evidence of base rate neglect (putting too little weight on the prior) as well as a representativeness bias (over-updating when sample draws match the population proportion) relative to Bayes' rule. Camerer (1987) found updating was well approximated by Bayes' rule, however found some evidence of representativeness that was less pronounced with more experienced subjects. Grether (1992) found evidence of conservatism, the opposite bias of base rate neglect, in addition to representativeness.

events such as being diagnosed with Alzheimer's disease or being robbed. They found that individuals updated more in response to good news relative bad news.[3] Another unincentivized study by Wiswall and Zafar (2015) finds some evidence that college students revise beliefs more when they receive information that average future earnings are greater than expected, relative to receiving information that earnings are less than expected.

As it is typically not possible to financially incentivize the elicitation of future life events, economists have turned to study belief updating about value relevant events in the laboratory. Crucially, these studies have differed from the early work, not only in context, but also in analysis. As there is little theoretical motivation for observing asymmetric updating in non-value relevant contexts, early work did not examine the relative weight placed on affirmative versus negative signals. With value relevant contexts, this changed. Importantly, this has implications for the comparability of results, as the wealth of analysis from earlier work cannot be brought to bear on more recent studies. This feature of the existing literature presents a clear rationale for the examination of robust counterfactual settings, an advantage of the current paper.

Previous and contemporaneous work on belief updating in value relevant contexts has focused on a particular context of interest, along with one counterfactual. Most relevant is Mobius et al. (2014), who pair an experiment with a theoretical model of optimally biased Bayesian updating in the context of ego relevant events. In the experiment they examine how individuals update beliefs about scoring in the top 50% on an intelligence test, using the same elicitation procedure as this paper. They find evidence that individuals update asymmetrically, over-weighting good signals relative to bad, and conservatively, updating too little in response to either type of signal. To provide evidence that these biases exist outside of ego relevant contexts they compare the results to a follow-up where a subset of the same subjects complete the updating task for a robot, and show that in the followup, both conservatism and asymmetry are reduced.

Regarding financial relevant events, Barron (2016) investigates updating beliefs when individuals have a financial stake in the outcome of drawing balls from two urns. His experiment complements this paper, focusing on one type of event while exogenously varying different values priors, as opposed to the across event variation that is the focus here. Barron (2016) does not find evidence of asymmetric over-weighting of good news when subjects could receive this "urn bonus", consistent with the results contained in this paper, that financial stakes do not lead to differential deviations in updating behavior. In contrast to the current paper financial stakes are substantially smaller (10 GBP rather than $80), and the

[3] Of note is that there is recent work which casts doubt on some of the evidence for asymmetric updating in these types of studies within psychology and neuroscience, see Shah et al. (2016).

focus precludes examining different types of events, including those that are ego relevant.

A number of other recent studies which have focused exclusively on ego relevant contexts are noteworthy. Buser et al. (2016) examine updating about performance on ego relevant tasks, with a structure similar to Mobius et al. (2014) and the present paper. Their focus is on examining heterogeneity in deviations from Bayes' rule by focusing primarily on within subject updating patterns. They find evidence of conservatism, but do not find overall asymmetry. They find evidence that conservatism is a stable trait of individual updating decisions, while asymmetry is not. As their results focus on within individual differences in updating behavior across ego relevant tasks, they do not examine updating behavior under an ego irrelevant control. In contrast, the focus of this paper is in examining updating across non-similar contexts, to provide evidence of whether biases in information processing are different for ego relevant or financial relevant contexts.

Eil and Rao (2011) study information processing about relative intelligence and beauty. Their results are broadly consistent with Mobius et al. (2014) with respect to asymmetric updating, however they do not find the same degree of conservatism.[4] Ertac (2011) also examines updating behavior regarding one's intelligence. Interestingly, she finds the opposite asymmetry, that bad news is weighted more than good news, similar to the results of this paper, though the conclusions differ starkly.

A common challenge inherent to all these papers involves the construction of an appropriate counterfactual updating task. Clearly, evidence of biased updating in value relevant contexts, does not imply the bias is specific to these contexts. Yet, earlier work has shown that individuals may update differently for different base rates or priors (e.g., base rate neglect, Grether (1980)), or for different distributions of received signals (e.g., representativeness, Grether (1992)). Even the objective versus subjective nature of the event may affect updating behavior, as this paper will demonstrate. Among the studies cited above, the majority have one counterfactual updating task, however only Barron (2016) (financially relevant) and Ertac (2011) (ego relevant) examine a counterfactual updating task that is robust to the above concerns, and in the case of Ertac (2011) only for a small subsample. In contrast, the current paper contains a more comprehensive set of updating decisions than previous or contemporaneous studies of belief updating. This broad set of decisions provides a natural set of counterfactuals that can be used to evaluate behavior. The results demon-

[4]Grossman and Owens (2012) conduct an experiment where absolute rather than relative performance is of interest to individuals. They do not observe the same biases in information processing of the studies on relative performance, however they do find that individuals hold overconfident priors about their absolute score on a quiz. Clark and Friesen (2009) find little evidence of overconfidence in a related study. In their study individuals revise beliefs after having experience with the task, as opposed to receiving a signal about their performance, making the conclusions difficult to compare with studies of biased updating.

strate that observing a broader set of decisions alters the interpretation of deviations from Bayes' rule in important ways.

Overall, there is mixed evidence on whether individuals update in a biased manner that over-weights good news relative to bad news. Of note is that the signal structure in both Eil and Rao (2011) and Ertac (2011) is different from Mobius et al. (2014), Buser et al. (2016), Barron (2016), and the current paper, which impedes direct comparability. In the results section, I return to a discussion of signal structure and surmise that some of the contradictory results in the literature may be related to these differences in signal structure.

3 Experimental Design

The experiment was conducted at New York University, at the Center for Experimental and Social Science (CESS). Recruitment was done via the CESS online system, where undergraduate students are notified by email when an experiment is scheduled, and have the opportunity to sign up. No restrictions were placed on participation, however subjects could only participate in the experiment once. A total of 326 subjects participated, in 32 different sessions for an average of 10 subjects per session.[5] The average subject payment was \$24.96 for approximately 75 minutes including a \$10 showup fee. The experimental data is also studied in Coutts (2015), there with the aim of distinguishing models of belief bias. That paper focuses exclusively on prior formation, and does not examine updating behavior.

Individuals in the experiment faced four different binary events and a sequence of four incentive compatible belief elicitations for each event. First, their prior beliefs about the probability of the event were elicited. Next they received a binary signal, regarding whether the event had occurred. This signal was true with two-thirds probability, and false with one-third probability. After receiving this signal their beliefs were again elicited, and the same process was repeated two more times.

One concern was that the elicitation procedure or the sequence of signals might be confusing to some subjects. These features of the experiment were presented making use of intuitive explanations that aid subject understanding, following Grether (1980) who recognized the importance of subject understanding and experimental credibility. A large component of the experiment consisted of intuitive explanations and examples, as well as practice with all of the experiment's components before the actual experiment began. A questionnaire was administered to every participant after completing the experiment that included an open ended question to address any issues of the experiment subjects found confusing or unclear.

[5] Because of a technical failure, one session resulted in data for only one event. 318 subjects participated in all four events.

This, as well as verbal feedback, suggested that subjects had a good understanding of the various procedures and components of the experiment.

3.1 Belief Elicitation

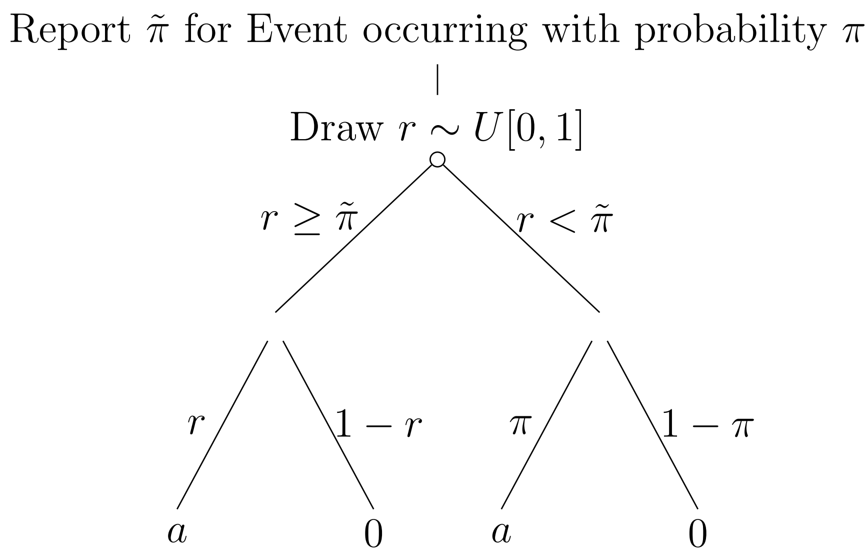
To elicit beliefs about the binary events I use the method of Karni (2009), utilized also by Grether (1992), Holt and Smith (2009), and Mobius et al. (2014), which I henceforth call the lottery method.[6] Incentive compatibility only requires that individuals exhibit probabilistic sophistication, see Machina and Schmeidler (1992), which loosely speaking requires that individuals are able to assign subjective probabilities to events and subsequently make choices given these probabilities. It does not require that individuals maximize expected utility, nor does it require assumptions on risk preferences.[7] The method involves the possibility of earning a lump sum payment a, and it's implementation is presented in Figure 1.

Subjects are asked to report a probability $\tilde{\pi}$ that makes them indifferent between participating in a lottery that pays a with probability $\tilde{\pi}$ and 0 otherwise, or participating in a lottery that pays a whenever the event of interest occurs. After indicating $\tilde{\pi}$ the computer draws a random number r distributed uniformly from 0 to 100. If $r \geq \tilde{\pi}$, a subject participates in the lottery that pays a with probability r. If $r < \tilde{\pi}$ the subject faces the lottery that pays a when the event in fact occurs ("event lottery"). When r takes on discrete values, this mechanism is equivalent to a 101 item choice list that requires a choice between the "event lottery" and an objective lottery which pays a with probability in the set of integers from 0 to 100, with one of the choices selected at random.

[6] See Karni (2009) for a more detailed description of the lottery method, though the method itself has been described in a number of earlier papers. See Schlag et al. (2014) for details about earlier descriptions of this mechanism. The mechanism is also referred to in various papers as the "crossover method", "matching probabilities", and "reservation probabilities".

[7] While this method is a variant of the Becker-Degroot-Marschak (BDM) mechanism, it is important to note that it is not subject to the critique of Karni et al. (1987), as the mechanism here identifies parallel probabilities that lead to indifference between two lotteries; it does not inquire about the monetary certainty equivalent of a lottery. For further discussion see Healy (2016).

Figure 1: The Lottery Method



Lottery method: First the individual submits a probability report $\tilde{\pi}$ for a binary event that occurs with true probability π. Next a random number r is drawn between $[0,1]$ using the uniform distribution. If $r \geq \tilde{\pi}$ the individual plays a lottery that pays out a with probability r, and 0 otherwise. If $r < \tilde{\pi}$ the individual is paid an accuracy payment $a > 0$ if the event occurs, and 0 if it does not occur.

In the experiment a is either low (\$3), medium (\$10), or high (\$20), randomized at the session level. For an intuitive way to think of this mechanism, imagine that an individual must tell a trusted friend to make a choice on their behalf in a private room, between the "event lottery" (which pays a if the event occurs) and a different "objective" lottery which pays a with some given probability. The friend doesn't know the probability of the objective lottery in advance, and so the individual must tell the friend at exactly which point they would prefer the objective to the event lottery.

In order to facilitate subjects understanding the lottery method, the experiment made use of a graphical interpretation meant to improve intuition.[8] Holt and Smith (2009), who utilized the same elicitation procedure, conducted "pencil and paper" experiments to give individuals practice with the method, and were on hand to answer any questions.[9]

In this experiment subjects were introduced to a gumball machine, that had 100 total

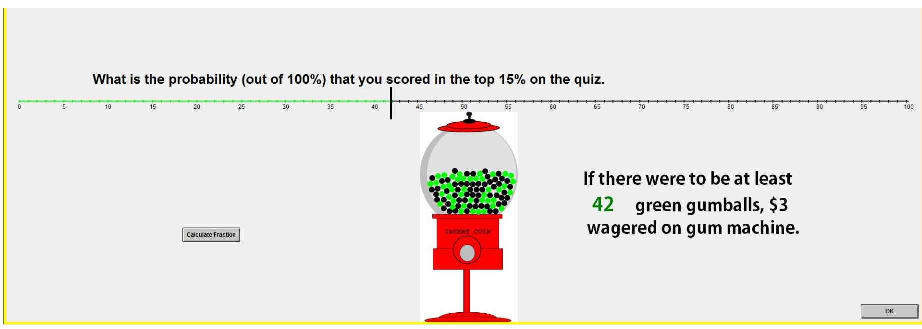
[8]Subject understanding of methods of belief elicitation has been a concern for many experimental economists. Schlag et al. (2014) provides a nice overview of the literature discussing the effects of complexity of elicitation procedures on responses. Schotter and Trevino (2014) is a good overview more generally on eliciting beliefs in laboratory settings.

[9]Mobius et al. (2014) conducted their experiment outside of the lab, via social networking website Facebook. As such providing real time feedback was not possible, however their design contained many intuitive elements meant to ease understanding.

gumballs, that were either black or green in color. This represented a lottery with the probability of success equal to the number of green gumballs out of 100. One gumball would be drawn from the machine, at random. Subjects were told that the computer had a list of 101 gumball machines, each with a different proportion of green gumballs, and that one of these would be randomly selected, i.e. the discrete uniform distribution.

They were then asked to indicate on a graphical slider, exactly what point (number of green gumballs) they would prefer to base their potential earnings on the "gumball lottery" instead of the "event lottery", which paid off if the event of interest occurred. Figure 2 provides screenshots of the gumball machine, as well as the slider subjects had to move. Subjects were given significant practice with the slider, with non-paid practice events, before the primary paid experiment began.

Figure 2: Screenshot from the Experiment: Slider



The gumball machine was used to provide an intuitive representation of the lottery method. Subjects would indicate on the slider the minimum threshold, i.e. the number of green gumballs there had to be in the machine before they would prefer to wager the accuracy payment (here $3) on the gumball machine rather than the event, "you scored in the top 15% on the quiz". The proportion of green gumballs in the picture adjusted as the slider was moved.

3.2 Events

There are two main sources of variation in the experiment, the events themselves, and the financial stakes. There are four events in total, which were presented to subjects in random order.[10] Of these four events, two involve rolling dice, and their probabilities can be objectively calculated.[11] The outcome of these events was determined by chance, and individuals could not affect these outcomes.

[10] One of the events (easy dice) was fixed as the final event. The other three events were randomly ordered at the session level. Updating behavior does not differ by order.

[11] It is difficult to find a rigorous definition of what makes a probabilistic process "objective". A good starting point is Gilboa and Schmeidler (2001), who provide a definition that agrees with the usage in this paper.

The other two events are subjective, and were based upon tasks that individuals had completed prior to the beginning of the experiment. Most relevant to previous studies of asymmetric updating is an event where individuals had to estimate the probability they scored in the top 15% on a 5 minute, ego relevant quiz. Percentiles in the quiz were generated by comparing scores to 40 individuals who took the quiz in previous piloting, which was known to students. The quiz emphasized multiple choice questions on math and English skills, similar to standardized college entry tests in the USA.[12] In order to generate a sensible control group, every individual in the experiment had a 30% chance of being selected to estimate the performance of a randomly selected anonymous partner in the room, rather than their own performance. 95 out of 318 subjects were thus randomly selected for this control, independent of any other treatments in the experiment.

The fourth event was whether the individual correctly answered a question about what the weather (mean temperature) was on a randomly selected day in New York City in the previous calendar year. This question is not objective in the sense of the dice questions, but it also does not appear to involve skill or ability.[13] Figure 3 summarizes the four events that all individuals faced.

One important feature of these events are the relatively low probabilities, due to budget constraints. Empirically the events occurred on average 15% of the time, with the implication that there will be significantly more negative than affirmative signals. In the data, 63% of the signals are indeed negative, which, as the results sections will discuss, may affect updating behavior. Importantly, while the probability of events were low, priors were significantly higher.[14] The average prior is 36%, and over one-third of priors are greater than or equal to 50%, which facilitates comparisons with other studies.

3.3 Stakes

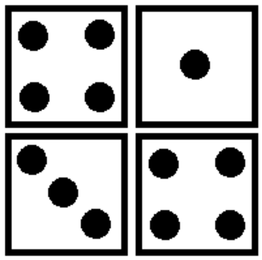
The other key source of variation in the experiment involved varying the financial stakes. For each event, individuals had a 50% chance of receiving an additional \$80 if that event occurred. This financial stake was made salient as subjects physically drew a token from a bag, that was labelled either \$80 or \$0. If they drew a \$0 token, they knew that they would have no financial stake in the event, they could only potentially earn an accuracy payment

[12]This quiz was taken before individuals made any choices, and before they had any knowledge of the four events. Additionally subjects were told truthfully, that performing better on the test would lead to higher expected payments in the experiment.

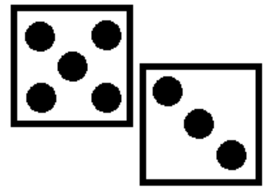
[13]Individuals were correct if the true temperature was within plus or minus 5 degrees Fahrenheit of their estimate. In fact, there is no correlation between beliefs about getting this question correct and actually getting the question correct.

[14]See Coutts (2015) for a discussion of this phenomenon.

Figure 3: Description of Events

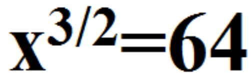


(a) Hard Dice: The computer rolls four dice. Event occurs when exactly two out of those four dice was a specified number (e.g. 4). The probability of this is $\binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 = \frac{150}{1296}$ or approximately 11.57%.



(b) Easy Dice: The computer rolls two dice. Event occurs when two different specified numbers were the only numbers to come up (e.g. 5-3, 3-5, 3-3, 5-5). The probability of this is $\frac{4}{36}$ or approximately 11.11%.

(c) Weather: Event occurs if the individual correctly estimated the average temperature on a specified random day in NYC in the previous year (2013), +/- 5 deg F. In the sample, 25.77% of subjects were in the correct range.



(d) Quiz: Event occurs if the individual scored in the top 15% on an ego relevant multiple choice quiz (self). For a subset of participants the event pertained to a random partner's performance instead of their own (other). Percentiles were generated in comparison to 40 pilot quiz-takers.

for their belief report. If they drew a $80 token, they knew that they could potentially earn $80, if the event in fact occurred, and they were selected for payment for that specific event.

In this latter case, they could still potentially earn the accuracy payment. However the design ensured independence between the financial stake and accuracy payment, to ensure incentive compatibility by eliminating hedging opportunities in a manner similar to Blanco et al. (2010). At the end of the experiment *only* the financial stake *or* the accuracy payment would be paid, chosen at random. Exact details of this procedure can be found in Appendix 6.1.

3.4 Signals (News)

News comes in the form of noisy, binary signals, after the first elicitation (prior). Signals were explained with the aid of pictures of "gremlins". This procedure is very closely related to that of Mobius et al. (2014). In their experiment they use two fictional characters to announce the signals, "Joke Bob" and "Wise Bob" each chosen with 50% chance. Wise Bob truthfully reveals the state of the world, while Joke Bob randomly (50/50) announces a state of the world. In this way, the chances of a signal being true are $\frac{3}{4}$ in Mobius et al. (2014).

In this experiment, individuals were told that there were three gremlins that all knew whether the event had occurred. Two of the three gremlins always told the truth, while one gremlin always lied. The subjects were then told one gremlin had been randomly selected, and that gremlin either provided them an affirmative signal (the event had occurred), or a negative signal (the event had not occurred).

In this way signals were true with probability $\frac{2}{3}$ and false with probability $\frac{1}{3}$, a fact that was emphasized to participants. After receiving the signal, a subject had his or her updated belief elicited, again using the lottery method. Subjects were given three independent signals (knowing the structure in advance) and were informed that the gremlins were drawn with replacement in order to maintain a constant probability of $\frac{2}{3}$ that the signal was true.

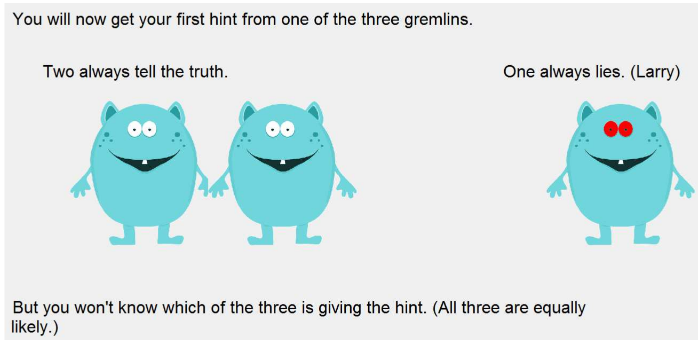
In total subjects had their beliefs elicited four times for each event: one prior elicitation, and three posterior. Since there were four events, every subject had their beliefs elicited 16 times in the experiment, of which 12 correspond to updated beliefs. One of the elicitation rounds was randomly selected at the end for payment.[15]

Figure 4 depicts screenshots from the experiment that showcase the use of "gremlins" as graphical aids.[16] In addition to these visual aids, individuals also had practice with receiving signals and updating beliefs with non-paid practice events, before the paid experiment began.

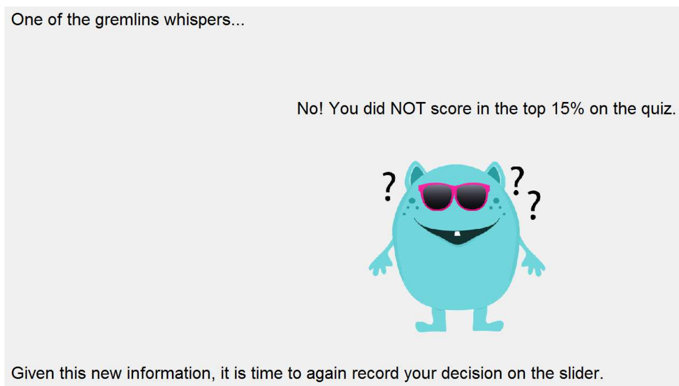
[15]As discussed in Appendix 6.1, in fact there was a probability subjects would end up in a *prize state*, which would result in their payoff depending on winning this prize, rather than the elicitation procedure.

[16]Images are from www.mycutegraphics.com, by author Laura Strickland.

Figure 4: Screenshots from the Experiment: Signals



(a) Screenshot introducing signals.



(b) Screenshot of a received (negative) signal.

3.5 Classifying Good/Bad News (Value Relevant) from Just News (Non Value Relevant)

To summarize the experimental treatments, accuracy payments of $a \in \{3, 10, 20\}$ were randomized at the session level, while the financial stakes of \$0 or \$80 were randomized at the individual-event level. Additionally, for the quiz event, 30% of subjects were randomly allocated (independent of the other treatments) to a control treatment where they were asked to update about another randomly selected individual's performance.

Thus, for an event that a given subject was allocated a \$80 financial stake in the event, information about whether the event occurred will contain good or bad news. Information which indicates such an event is more likely, corresponds to an increase in the expected probability of earning that \$80 stake, and vice-versa.

For those events that a given subject was allocated a \$0 stake, whether information/news is good or bad depends on the event itself. To the extent that individuals gain utility from believing they have high ability, the quiz event (estimating own performance) involves a personal, non-pecuniary stake. Thus for this quiz event, binary signals about performance will contain either good news (they are in the top 15%), or bad news (they are not in the top 15%).

These contexts are in contrast with non value relevant events where there are no personal stakes (the two dice events and the weather event), *and* those in which subjects held a $0 (rather than $80) financial stake in the outcome. Here signals contain information about outcomes, but these outcomes are irrelevant to individual well being. In other words, news is simply news.[17]

Of 1280 event observations, 634 involved a financial stake of $80, which means signals regarding these events are categorized as value relevant. This means 646 events had no financial stake.[18] However, of these 646 events, 115 were the quiz (self) event which is a personal stake. Thus altogether, I classify 749 events as potentially relating to good/bad news (value relevant), and 531 events as just news (non value relevant).

4 Results

4.1 Overview of Pooled Results

As a first pass at examining updating behavior I plot reported subject beliefs and compare these to posterior beliefs that would have resulted if subjects updated using Bayes' rule. Following the discussion in the previous section, I split the sample into events where subjects have a financial or personal stake in the outcome, and events where subjects have no such stake.

The former include all events where subjects stand to earn $80 if the event occurs, as well as the event that involved whether they scored in the top 15% on the quiz. For these events, information in the form of noisy signals is valenced. The latter events do not involve individual ability and subjects have no additional financial stake in the outcome of the event. Thus, signals provide new information about the outcomes of these events, however this information is neither good nor bad. All belief reports are financially incentivized using the lottery method described above.

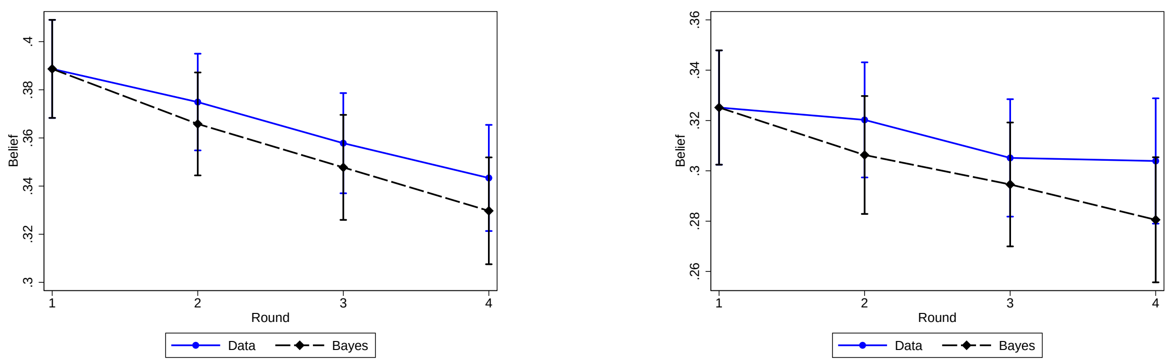
Figure 5 plots the average reported belief for the prior, as well as the belief after receiving each signal. Included in the same figure is the path that beliefs would take if individuals were

[17]As noted before, accuracy payments are a form of financial stake. However, ex-ante, an individual has no reason to care whether an event occurs or not, as long as their belief report is accurate. It may be argued that the weather event involves a personal stake, if individuals derive utility from correctly estimating historical temperatures. I do not find this likely, but regardless, none of the results in this paper hinge on the inclusion of the weather event.

[18]The reason they are not equal is that some sessions had odd numbers of subjects, and the physical drawing of $0 or $80 always involved an equal proportion of both stakes.

perfect Bayesians, given the subject's first reported belief. That is, the Bayesian benchmark takes only the prior from the data and then calculates how beliefs would evolve given the signals that subjects actually received.

Figure 5: Evolution of Beliefs



(a) Financial/Personal Stake: Good/Bad News **(b)** No Stake: Just News

The path of beliefs starting from the prior (round 1) and after each sequential signal (rounds 2 through 4). **(a)**: Quiz (self) event and all instances where subjects could earn \$80 if the event occurred. **(b)**: All other instances. Average individual responses are the blue solid line, the Bayesian benchmark is marked as the black dashed line. Bayesian benchmark takes prior beliefs, and subsequently uses Bayes' rule to update beliefs. Error bands represent 95% confidence intervals. **(a)** $N = 749$ **(b)** $N = 531$ observations per round.

From Figure 5 one can see that there are slight deviations from Bayes' rule for both subsamples, as individual's update slightly more conservatively. However the difference between reported posteriors and Bayesian posteriors is not significant at any conventional level. There are slight differences in prior belief formation across the two groups, part of which is accounted for by the presence of the quiz (self) event. In fact, priors are biased upwards for all events, as detailed in Coutts (2015).

There do not appear to be any substantive differences in patterns of updating across the two subsamples. Thus from an initial look at the data, updating does not appear to differ when news is good or bad, compared to when news is simply news.[19] For the aggregate data the correlation between empirical posterior beliefs and posterior beliefs calculated using Bayes' rule is 0.89, higher than that found in Mobius et al. (2014). Pooling across the three updating rounds, the average posterior is 33.8%, while the average Bayesian posterior, calculated using subject priors, is 32.5%. These are remarkably similar, though it is noteworthy

[19]In fact, in Appendix 6.2 I present these figures for all events, all financial stake, and all accuracy payment conditions separately. From those figures, one can see that there are some slight differences in updating behavior across events, but financial stakes or accuracy payments do not appear to alter updating behavior. Examining only the event about performance on the quiz, unlike previous studies, updating about own performance does not appear to deviate much from the Bayesian prediction.

that the difference is significantly different from zero at the 10% level. Interestingly, despite the high correlation between actual posteriors and Bayesian posteriors, the aggregate data includes a large number of non-updates. 41% of reported posteriors are identical to reported priors and only 9% of subjects update in every round.

To summarize thus far, updating appears to be well approximated by Bayes' rule, despite a large number of non-updates. However, looking only at posterior beliefs does not provide detailed information about how individuals react to signals or news about outcomes, and may be affected by the fact that negative signals are more prevalent than affirmative signals. In the next section I use a flexible empirical framework to examine how individuals respond to both affirmative or negative signals, which, depending on the event and stake conditions, may be interpreted as good or bad news, respectively. This permits a more rigorous investigation into whether individuals will update asymmetrically, when the outcomes of events are either ego relevant or financially relevant, as has been found in previous literature.

4.2 Information Processing: Framework

I now investigate the updating process in more detail, to examine whether the above analysis masks asymmetries in updating behavior. For the analysis in this section I follow Mobius et al. (2014) and use a flexible model of updating that retains the structure of Bayes' rule, but allows for the possibility that individuals place different weight on the prior, and/or affirmative or negative signals. The model is a variant of that used originally by Grether (1980). Bayes' rule can be written in the following form, considering only binary signals, $s_t = i \in \{0, 1\}$, and letting $\hat{\mu}_t$ be the belief at time t:

$$logit(\hat{\mu}_t) = logit(\hat{\mu}_{t-1}) + I(s_t = 1)\gamma_1 + I(s_t = 0)\gamma_0 \tag{1}$$

where $I(s_t = i)$ is the indicator function when the signal $s_t = i \in \{0, 1\}$, and γ_i is the log likelihood ratio of the signal being $i \in \{0, 1\}$, where $\gamma_0 = \ln \frac{1}{2} = -\ln 2$ and $\gamma_1 = \ln 2$ given the signal strength of $\frac{2}{3}$. The empirical model nests this Bayesian benchmark in the following way:

$$logit(\hat{\mu}_{it}) = \delta logit(\hat{\mu}_{i,t-1}) + \beta_1 I(s_{it} = 1)\gamma_1 + \beta_0 I(s_{it} = 0)\gamma_0 + \epsilon_{it} \tag{2}$$

δ captures the weight placed on the prior belief. β_0 and β_1 capture responsiveness to either negative or affirmative signals respectively. Bayes' rule is a special case of this model when $\delta = \beta_0 = \beta_1 = 1$. In the context of the experiment, $s_{it} = 1$ corresponds to a signal that YES the event had occurred, while $s_{it} = 0$ corresponds to a signal that NO the event had not occurred. Since $I(s_{it} = 0) + I(s_{it} = 1) = 1$ there is no constant term. ϵ_{it} captures non-systematic errors.

Additionally, as described in Mobius et al. (2014), Bayes' rule also satisfies three additional properties: invariance, sufficiency, and stability. When $\delta = 1$, the updating process is said to satisfy invariance, i.e. the change in logit beliefs depends only on past signals. Sufficiency requires that after controlling for prior beliefs, lagged information does not significantly predict posterior beliefs. Finally, stability requires that the structure of updating is stable from round to round. These properties of the data are investigated in Appendix 6.4. Overall, the pooled data do *not* support these properties, though the magnitude of deviations is relatively small.[20]

Table 1 presents the specification in Equation 2 for the aggregate data, as well as the two subsamples introduced earlier, pooling across all updating rounds. Note that asterisk's indicate significantly different from the Bayesian benchmark prediction of one, *not* significantly different from zero. Standard errors are clustered at the individual level. In the primary sample I do not include posterior beliefs that were updated in the opposite direction that Bayes' rule predicts, which amounts to dropping 4.8% of observations.[21] I include all other subjects, including those who never update their beliefs.[22]

Table 1 provides the finer details of updating behavior that Figure 5 is unable to capture. What is first apparent is that updating behavior deviates from the strict Bayesian prediction that all coefficients are equal to 1. There is substantial conservatism in response to both affirmative and negative signals, as indicated by coefficients less than one for β_1 and β_0. Of note is that the degree of conservatism is less than that found in studies by Mobius et al. (2014) and Buser et al. (2016).

An asymmetric bias is also clearly present across Table 1, with negative signals receiving more weight than affirmative. The null hypothesis that affirmative signals receive the same

[20]One valid concern regarding the OLS analysis is in using priors as a dependent variable. Since priors are lagged posteriors, this creates a potential issue if there is substantial heterogeneity in response to signals, which could lead to upwardly biased estimates of δ, see Mobius et al. (2014). Instrumenting with higher order lagged beliefs or lagged Bayesian beliefs is possible, however such techniques do not alter the results reported. Recovering unbiased estimates of δ is also not central to the results of this paper.

[21]4.8% is less than the approximately 10% in Mobius et al. (2014) and Buser et al. (2016).

[22]Because Bayesian posteriors are never at the boundary for intermediate priors, the framework is agnostic for priors or posteriors equal to 0 or 1. Following Mobius et al. (2014) and Buser et al. (2016), these observations are dropped, amounting to 6% of the sample. In Appendix Table 8 I examine potential implications of these sampling restrictions, following Grether (1992) and Holt and Smith (2009) by replacing boundary observations by 0.01 or 0.99 respectively. Intuitively, there is evidence that including boundary observations reduces conservatism, suggesting this may be important in understanding conservatism.

Table 1: Updating Beliefs for All Events

Regressor	Good/Bad News	Just News	All
δ	0.918***	0.907***	0.914***
	(0.012)	(0.014)	(0.009)
β_1	0.594***	0.576***	0.588***
	(0.040)	(0.051)	(0.034)
β_0	0.782***	0.812***	0.793***
	(0.043)	(0.051)	(0.038)
P-Value ($\delta = 1$)	0.0000	0.0000	0.0000
P-Value ($\beta_1 = 1$)	0.0000	0.0000	0.0000
P-Value ($\beta_0 = 1$)	0.0000	0.0003	0.0000
P-Value ($\beta_1 = \beta_0$)	0.0002	0.0003	0.0000
R^2	0.84	0.82	0.84
Observations	1950	1410	3360

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

weight as negative signals is rejected at the 1% level. Again, there do not appear to be any differences between beliefs about events where subjects have a personal or financial stake versus those where subjects have no stake, and I cannot reject equality of any of the three coefficients across the subsamples at conventional significance levels. Thus whether news is a source of good/bad information or simply neutral information, does not appear to alter how beliefs are updated.

Similar to the previous section, there are no significant differences in updating behavior across the financial stake conditions, including the varying payments for accuracy, all of which is presented in Appendix 6.3. Regarding the accuracy payments, this is relevant to studies of the effects of stakes on behavior in lab experiments, especially regarding belief elicitation. It suggests that paying subjects more for accurate beliefs may have little effect on belief updating. Combined with Coutts (2015), which showed that increasing accuracy payments can lead to more biased prior formation, the implications are that ideal incentive payments may be relatively low. In the next section I present the same analysis of updating behavior across events, to better understand whether and how these deviations differ across events.

4.3 Information Processing By Event

While the previous results showcased that posteriors are well approximated by those generated by Bayes' rule, some important deviations remain in the structure of updating. This section examines whether these deviations appear in specific events. Previous evidence has found that individuals update asymmetrically when provided information on their performance on a test, over-weighting good news relative to bad as in Mobius et al. (2014) and Eil and Rao (2011), or the opposite asymmetry as in Ertac (2011). I thus focus attention on the quiz (self) event: whether an individual believes they scored in the top 15% on an ego relevant quiz. The signal structure and elicitation procedure is comparable with Mobius et al. (2014) and Buser et al. (2016), who examined beliefs of subjects about scoring in the top 50%, rather than 15%.

Note that in studying updating behavior in one context, it is critical to have an appropriate counterfactual comparison. Deviations from Bayes' rule in a given context do provide evidence against Bayes' rule, but attribution of such deviations to a particular bias or context is only valid if one can rule out that these deviations occur in other contexts. Use of an adequate control group is standard for extrapolating that deviations do not occur across other contexts. This requires that the control group is suitably defined, and has adequate statistical power to rule out deviations of interest. Due to the difficulties in defining an appropriate, high powered control group, few studies in this literature are able to satisfy this requirement. As I will discuss further, defining a suitable control group for updating on the quiz is not necessarily straightforward.

As a control group for the quiz event in this experiment, 30% of subjects do not update about their own performance, but instead must update about the performance of a randomly selected anonymous individual in the lab. This is an intuitive control, yet one issue is that observed prior beliefs about one's own performance tend to be greater than those about another subject's performance. This is problematic, as in the Appendix, in Figure 9, I present some evidence that updating differs for different values of the prior.

In Mobius et al. (2014), the control group comes from a follow-up with subjects using randomly generated performance of a computer matched to the subject's (stated) expected performance. This circumvents the issue of differently sized priors, however it creates a new issue, that subjects might not update similarly for objectively versus subjectively given priors.

Despite the difficulties and potential doubts regarding selection and validity of the control group, the existence of evidence across contexts of this experiment provide additional comparison groups. There is enough variation across the different events and financial treatments such that differences in updating behavior in one context that do not appear in any other would be relatively strong evidence that such patterns are indeed specific to the context. On the other hand, similar differences in updating across contexts would suggest that deviations from Bayes' rule may reflect more general updating biases, rather than being unique to a specific context. For instance, Grether (1992) found that when studying updating about neutral events, there were differences in updating behavior even across contexts where the underlying inference tasks were the same.

Table 2: Updating Beliefs Within Events

Regressor	Easy Dice	Hard Dice	Weather	Quiz (S)	Quiz (O)
δ	0.894***	0.872***	0.928***	0.952**	0.912**
	(0.028)	(0.022)	(0.024)	(0.022)	(0.035)
β_1	0.476***	0.404***	0.684***	0.590***	0.709**
	(0.090)	(0.062)	(0.061)	(0.054)	(0.118)
β_0	0.821*	0.886	0.818***	0.834***	0.732***
	(0.099)	(0.080)	(0.047)	(0.060)	(0.099)
P-Value ($\delta = 1$)	0.0002	0.0000	0.0033	0.0259	0.0130
P-Value ($\beta_1 = 1$)	0.0000	0.0000	0.0000	0.0000	0.0157
P-Value ($\beta_0 = 1$)	0.0719	0.1538	0.0001	0.0063	0.0082
P-Value ($\beta_1 = \beta_0$)	0.0398	0.0000	0.0663	0.0005	0.8942
R^2	0.73	0.77	0.75	0.84	0.79
Observations	836	841	871	565	247

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

Table 2 presents the results of all rounds of updating corresponding to each of the four different events, splitting the quiz event self or other performance.[23] Examining updating across domains, there are some differences in updating behavior, yet these differences do not appear to fit a consistent pattern.

The most suggestive result is found in Columns 4 and 5, which respectively compare updating behavior for the quiz event when it involves one's own performance versus another subject's performance. Regarding one's own performance, there is asymmetric underweighting of good news relative to bad news, significant at the 1% level. Yet for the comparison group, other performance, there is no significant asymmetry in updating.

This result is the opposite asymmetry of that found in Mobius et al. (2014), but is consistent with evidence from Ertac (2011). Yet, this result is greatly undermined when comparing estimates across other events. Even larger asymmetries are found for the two dice events, while a slightly smaller asymmetry can be seen for the weather event. Since these

[23]These estimates are not substantively altered using the same sampling procedure as Mobius et al. (2014).

events involve outcomes that subjects have no personal stake in (and there is no difference for financial stakes), it is clear that the differences in asymmetric updating between own and other performance, may be driven by factors that affect information processing more broadly, i.e. general cognitive biases, rather than a specific psychological bias. Importantly, these conclusions do not hinge on particular values of the prior, as even moderate values of the prior are associated with the same asymmetry.

Thus, two conclusions emerge from the data examining updating across domains. The first, is that I do not find evidence of over-weighting good news when it pertains to information about one's intelligence, that has been found in some, but not all previous studies. The second finding is that without the comparison across domains, comparing the quiz (self) event with the control (quiz other) would lead to very different conclusions. This is because there is a sizeable asymmetry (over-weighting bad news) that is significant at the 1% level for updating beliefs about one's performance, whereas for another's performance the difference is smaller and not significant. I can also reject at the 1% level that the asymmetry observed for one's own performance is equal to the point estimate of the asymmetry for the comparison group.

Yet, I am unable to reject the hypothesis that the response to signals for the quiz (self) event is the same as all other events pooled, whether aggregated or partitioned into good/bad news versus just news subsamples. The evidence suggests that the observed deviations from Bayes' rule are not generated by differential responses to good or bad news about performance, as they are present across other events that don't involve performance nor do they involve good or bad news. This has important implications for the empirical and theoretical study of biased belief updating, particularly the emphasis on asymmetric updating in previous work. It suggests caution in interpreting differences in updating patterns within a specific event as evidence of ego relevant psychological bias, without examining for the presence of cognitive bias in other neutral contexts.

4.4 Understanding Asymmetry

The previous section demonstrated that subjects update asymmetrically across most contexts, including those that relate to quiz performance, as well as for neutral events that involve dice rolls which are ego irrelevant. This result fits in a literature that has found both evidence of positive asymmetry regarding good news versus bad news, as in Mobius et al. (2014) and Eil and Rao (2011), but also negative asymmetry found by Ertac (2011). In this section I explore a possible unifying explanation for these mixed results: signal structure.

One difference from the present paper and that of Mobius et al. (2014) and Eil and Rao

(2011) is that the distributions of received signals are biased towards negative signals, due to binary events with average probabilities less than one-half. In this section I seek to understand whether there could be differences in how subjects update across different distributions of received signals. For example, subjects may receive three affirmative signals, two affirmative and one negative, one affirmative and two negative, or three negative signals. Table 3 examines how subjects update given the sequence of signals they faced in the experiment.[24]

Contrary to the Bayesian prediction, there are clear differences in updating given different sequences of signals. Examining Table 3, there is a pattern of underweighting signals that are received less often. This also means that when subjects receive two affirmative and one negative signal, the asymmetry is reversed, and in fact affirmative signals receive more weight than negative signals. This is important, because given the unlikely nature of many of the events in this experiment, the distribution of signals is more heavily weighted towards negative signals. The implication is that the observed negative asymmetry in the data may in part be accounted for by the relatively large number of negative signals.

These patterns are suggestive that differences in signal structure may be able to account for the stark differences observed in updating behavior in recent work by Mobius et al. (2014), Eil and Rao (2011), and Ertac (2011). A major difference between all three studies is in the signal structure. In Mobius et al. (2014), the structure is similar to this paper with a signal strength of $\frac{3}{4}$ rather than $\frac{2}{3}$. In Eil and Rao (2011), subjects have their beliefs elicited about what decile they believe their performance lies in, and receive a perfectly informative signal about whether they are ranked higher or lower than a randomly chosen subject. In both Mobius et al. (2014) and Eil and Rao (2011), received signals were reasonably balanced between affirmative and negative signals.

In Ertac (2011), subjects must update about the probability their performance falls in the top, middle, or bottom. Rather than receiving signals of predetermined strength as in Mobius et al. (2014) and the current paper, they receive information that their performance was or was not in the top/bottom, which in some cases was fully revealing. While this type of signal structure makes comparison difficult, at a superficial level, as in the current paper, the types of signals of Ertac (2011) were also less likely to be affirmative: top (bottom) was less likely than not top (bottom). Nonetheless, because a signal of top (bottom) would completely reveal the state, it is not clear that the same asymmetry should persist when updating among the remaining possible states.

Across these papers there is also variation in the size of the prior as well as differences

[24] Note that for the Quiz (self) event, the distribution of signals faced may depend on ability as higher scoring individuals are likely to receive more affirmative signals. Excluding this event does not alter the results.

across events themselves, suggesting that opposing findings may not be unfounded. However, note that even when focusing on priors around 50% (common to many studies of belief updating) the results of this paper are unaltered. On the whole, the evidence in this paper suggests that some of the results across the literature may not be inconsistent if they are related to more general patterns of cognitive bias in updating beliefs across different information environments. An important direction for future work would be to determine how the signal structure itself affects information processing patterns more generally.

Table 3: Updating Beliefs By Distribution of Signals Received

Regressor	0 '+ Signals'	1 '+ Signal'	2 +' Signals'	3 '+ Signals'
δ	0.900***	0.903***	0.915***	0.982
	(0.013)	(0.014)	(0.029)	(0.068)
β_1		0.429***	0.804***	1.214
		(0.032)	(0.059)	(0.152)
β_0	0.819***	0.826***	0.556***	
	(0.043)	(0.056)	(0.105)	
P-Value ($\delta = 1$)	0.0000	0.0000	0.0032	0.7958
P-Value ($\beta_1 = 1$)	.	0.0000	0.0010	0.1634
P-Value ($\beta_0 = 1$)	0.0000	0.0020	0.0000	.
P-Value ($\beta_1 = \beta_0$)	.	0.0000	0.0353	.
R^2	0.82	0.83	0.78	0.78
Observations	1416	1435	441	68

Analysis uses OLS regression. K '+ Signals' refers to K affirmative signals, out of a possible maximum of 3. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

4.5 Conservatism, Ability, and Gender

In this section I discuss the substantial conservatism, largely driven by a failure to update beliefs after receiving a signal, and discuss its relationship with ability and gender. One important result is that observed conservatism is driven entirely by non-updates amounting to 41% of the entire sample. In the Appendix Table 9, I present the analysis of Table 1, but only for actively updated belief reports. From Table 9, one can see that conservatism disappears, as the estimated coefficient on the response to both affirmative and negative signals is greater than 1.

Critically, these non-updates are not simply driven by a small subset of conservative individuals: only 9% of subjects update in all 12 rounds, and the median subject updates in

7 of 12 rounds.[25]

In the Appendix, Table 10 presents some reduced form estimates of the factors that correlate with active updates. An important factor appears to be the size of the prior. A quadratic function of the prior belief predicts the prior that is most likely to lead to an update is 0.46, and is decreasing for priors towards either extreme of 0 or 1.

Looking across events, subjects update slightly less for the two objective dice events than for the other three events which are more subjective, though some of this can be accounted for with individual specific fixed effects. Additionally, subjects update more frequently in later updating rounds, the probability of an update is approximately 5% greater after receiving the second signal, and 14% greater after the third.

Examining Column 3 of Table 10, I also examine the relationship between active updating and the percentile rank on the quiz, coded from 0.01 to 1. Subjects who rank one standard deviation higher on the quiz are 5.7 percentage points more likely to update, significant at the 1% level. Interestingly, the positive association between ability and updating propensity is entirely driven by women. Men who rank poorly on the quiz update more frequently than equal ability women, but scoring higher on the quiz increases the propensity to update for women, the opposite of men.

Related, in Table 4 I follow the earlier framework to examine how gender and ability affect updating behavior more generally. I do this by allowing heterogeneous response to signals by gender and percentile rank.[26] The first column examines only interactions with gender, the second examines ability, while the third interacts the two. Similar to Mobius et al. (2014) and Ertac (2011) I find that women update more conservatively than men. There is however, however, no difference in asymmetry between men and women.

One interesting result is that unlike evidence from Mobius et al. (2014) and Buser et al. (2016) (see Barber and Odean (2001) for an earlier summary), women are not less confident about their performance than men on the quiz. Even controlling for ability, men are in fact slightly less confident than women, though the difference is not statistically significant at conventional levels.

Regarding ability, it appears that ability is related to both conservatism and asymmetry. In fact, subjects with higher ability put more weight on affirmative signals, but not on negative signals. The implication is that those at the top end of the ability distribution would no longer exhibit significant asymmetry, and would also be less conservative. This is in contrast to the results of Mobius et al. (2014) who found that neither conservatism nor

[25] Conservatism is correlated across events, within individuals, as found in Buser et al. (2016). 30% of the variation in non-updates can be explained by individual fixed effects.

[26] One could also interact gender and ability with the weight on prior beliefs, δ. I do not report these estimates, but interactions with δ are not significant.

asymmetry were significantly correlated with cognitive ability.

More broadly, the pattern of women updating more conservatively than men is present across both contexts where subjects have a personal or financial stake and contexts where there is no such stake. Thus while previous work from Ertac (2011) and Mobius et al. (2014) raised the important possibility that female conservatism may be related to self-confidence, such conservatism is equally present in domains which are irrelevant to personal qualities.[27] Moreover, when comparing the deviations from Bayes' rule of final posterior beliefs of men and women, women are only one-tenth of one percentage point further from Bayes' rule than men, a difference that is not statistically significant.

Table 4: Updating Beliefs by Ability and Gender

Regressor	Gender	Ability	Gender × Ability
δ	0.910***	0.914***	0.910***
	(0.010)	(0.009)	(0.010)
β_1	0.520***	0.433***	0.394***
	(0.047)	(0.076)	(0.103)
β_0	0.740***	0.776***	0.728***
	(0.048)	(0.071)	(0.097)
$\beta_1\times$ Male	0.141*		0.135
	(0.077)		(0.159)
$\beta_0\times$ Male	0.160**		0.198
	(0.065)		(0.142)
$\beta_1\times$ Percentile		0.315**	0.276
		(0.133)	(0.190)
$\beta_0\times$ Percentile		0.037	0.028
		(0.120)	(0.169)
$\beta_1\times$ Male × Percentile			−0.031
			(0.270)
$\beta_0\times$ Male × Percentile			−0.078
			(0.250)
R^2	0.84	0.84	0.84
Observations	3199	3360	3199

Analysis uses OLS regression. Difference is *significant from 0* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

[27] Importantly however, if gender differences persist across domains in updating, the same implications remain that high ability women may end up less confident than high ability men. An important direction for future research would be to better understand the source and implications of differences in information processing by gender.

5 Concluding Discussion

In this experiment I set out to examine whether people process information in a Bayesian way, in light of recent evidence from economics and neuroscience that individuals over-weight good news relative to bad news. In this experiment I study updating in contexts where subjects receive value relevant or non value relevant signals. This was achieved by randomly providing some subjects with a sizeable financial stake in certain events, as well as by examining updating about ego relevant contexts such as performance on an ego relevant quiz.

Unlike some of the claims and evidence of previous literature, I do not find evidence of asymmetric over-weighting of good vs bad news. In fact I find evidence of the opposite asymmetry, as in Ertac (2011), yet such patterns are similarly present comparing information processing across different contexts. Thus updating appears similar, whether it involves news regarding events that directly affect subjects' wellbeing, or whether it involves only information about neutral events. This result suggests that asymmetric updating found in previous studies is not a universal property of updating in ego relevant contexts. Further, it suggests that deviations from Bayes' rule may not be driven by psychological bias in response to good or bad news, but instead may reflect more general cognitive biases.

While these results differ from much of the previous literature, they do suggest a possible unifying feature. The distribution of signals received appears to affect how individuals update beliefs. Differences in the direction of the asymmetry, for example between, Mobius et al. (2014) and Eil and Rao (2011), and Ertac (2011), might be explained by differences in the type and content of information. High levels of conservatism appear more likely for more extreme values of the prior, and the direction of asymmetric updating appears related to the types of signals received. Moreover, these patterns in the data are present across all events and stake conditions, which suggests caution in attributing previously observed bias to one particular context. This is important, as contexts such as updating about intelligence have faced much greater scrutiny due to the psychological plausibility of asymmetric bias, compared with updating about neutral events where asymmetry is not expected, and therefore has not been studied.

If asymmetry and conservative patterns in updating represent how individuals process information more generally, this can be accounted for by appropriately defining control or comparison groups. Yet as this paper has discussed, such control groups can be difficult to define as subjects may update differently depending on (1) the prior, (2) the mix of signals received, and (3) whether the event is based on objective versus subjective uncertainty. To my knowledge, only Ertac (2011) designs an experiment with these factors in mind, however her control groups were varied at the session level, subsequently appropriate comparison is only possible for a small subsample.[28]

It is also important to point out differences in implementation between this experiment and that of previous work, particularly of Mobius et al. (2014). In Mobius et al. (2014) subjects participated in the experiment online, while in this experiment, subjects participated in an experimental lab. Further, it may be that differences being in the top 15% versus 50% played a role. While signal strength differed slightly, recent work by Ambuehl and Li (2015) suggests that subjects do not react strongly to signal precision. Future work is needed to understand the source of these differences.

Regarding gender differences in updating behavior my results suggest that, as found in earlier studies, there are important differences in updating beliefs by gender. The evidence suggests that these gender differences are not driven only by events where subjects have financial or personal stakes, but are present even in neutral contexts. Overall, while I do find that many predictions of the Bayesian model are rejected in the pooled data, these discrepancies are small in magnitude. The average posterior belief is less than one percentage point away from the posterior predicted by Bayes' rule. After three rounds of updating, subjects are less than two percentage points away from the posterior that would have resulted had they used Bayes' rule all the way through.

The results of this paper also relate to literature on the effect of stakes on behavior in experiments. Interestingly, paying subjects \$3, \$10, or \$20 for accurate belief reporting does not alter updating behavior. This has implications for budget-constrained experimental economists who are interested in incentivizing studies of belief updating, suggesting that paying greater amounts of money does not alter updating behavior.

Despite the large literature in psychology and economics discussing contexts where individuals process information differently from what Bayes' rule predicts, there have been few successful counter-theories that can match behavior in a variety of settings as well as Bayes' rule can. More recent evidence has suggested that domains involving personal characteristics are plausible contexts for deviations from Bayes' rule, particularly as a result of asymmetric information processing of good news relative to bad news.

This paper presents counter evidence to these studies, showing that updating may not be asymmetric in response to different types of signals for all types of ego boosting events. Such evidence is necessary to discipline existing and future theoretical work on updating behavior, in order to be better able to predict how individuals respond to information.

[28]In particular only 99/4640 observations can be matched with the most appropriate control group in her study. For non-ego settings, Barron (2016) analyzes a sensible counterfactual.

6 Appendix

6.1 Hedge Proof Design Details

Payoffs are determined in the following way, also described in Figure 6. To my knowledge this method was first utilized in Blanco et al. (2010).[29] In order to ensure that agents do not wish to hedge their probability reports, the world is partitioned into two disjoint states, the accuracy state and the prize state.[30] With probability 0.5 the individual is paid solely according to her reported belief $\tilde{\pi}$ about whether event E occurred using the incentive compatible lottery method to elicit beliefs with an accuracy payment of $a > 0$ (*accuracy state*).[31]

In the second state occurring with probability 0.5, the individual receives a guaranteed payment $\bar{a} \geq a$[32] and receives an additional \$80 if E occurs, but receives nothing extra if E does not occur (*prize state*). Her report of $\tilde{\pi}$ is no longer relevant in this prize state.

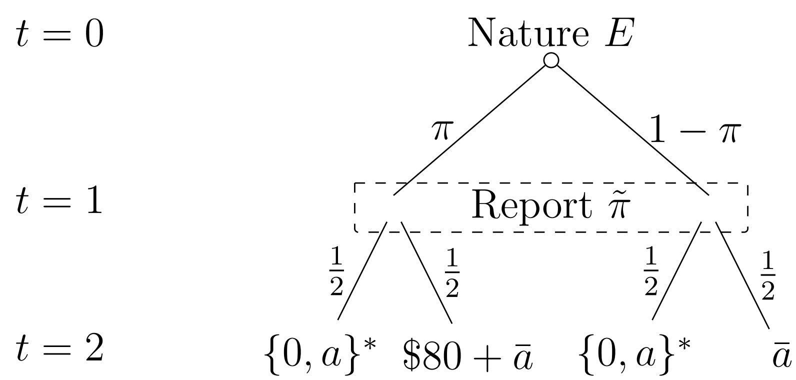
[29] It was also independently suggested to me by Christopher Woolnough, who I must credit for the design in this paper.

[30] Hedging will be present whenever utility is not linear, for example with a concave utility function and a positive stake in an event an individual would prefer to report a lower than truthful $\tilde{\pi}$, since this will smooth consumption over the different states of the world. Karni and Safra (1995) show that without this partition, no elicitation procedure exists that induces truthful reporting, a fact that is sometimes overlooked in the experimental literature; see Armantier and Treich (2013).

[31] To be clear, two types of hedging are of concern in this experiment. The first is hedging across accuracy and prize states, which is solved through partitioning. The second is hedging within the accuracy state, which is solved through use of the lottery method.

[32] The payment of $\bar{a}$ is to ensure that the prize state is always preferred to the accuracy state, required for an earlier theoretical extension, but not a necessary assumption for any of the analysis.

Figure 6: Illustration of Hedge Proof Design



*In the accuracy state the payoff is either 0 or a, depending on the reported belief $\tilde{\pi}$ and whether E occurred, according to the lottery method.

Nature determines outcome of binary event E. Individual submits report $\tilde{\pi}$ without knowing outcome of E, and payoff is determined according to the lottery method elicitation procedure.

6.2 Updating Patterns: By Event/Stake/Accuracy Payment

In this section I examine patterns in updating behavior for different events and financial stake conditions. Recall that the lump sum payment used for the lottery method was randomized at the session level, and was either \$3, \$10, or, \$20. The financial stake was randomized at the individual-event level, and was either \$0 or \$80 with 50% probability respectively. The financial stake was an amount of money that would be gifted to the subject if the event occurred and had been randomly selected for payment.

In Figure 7 I examine the analog to Figure 5, for each of the two financial stake conditions (\$0 and \$80), as well as each of the three accuracy payment conditions (\$3, \$10, \$20). While different values of the accuracy payment do not affect whether news is good or bad, note that having an \$80 stake in an event necessitates that signals contain either good or bad news.

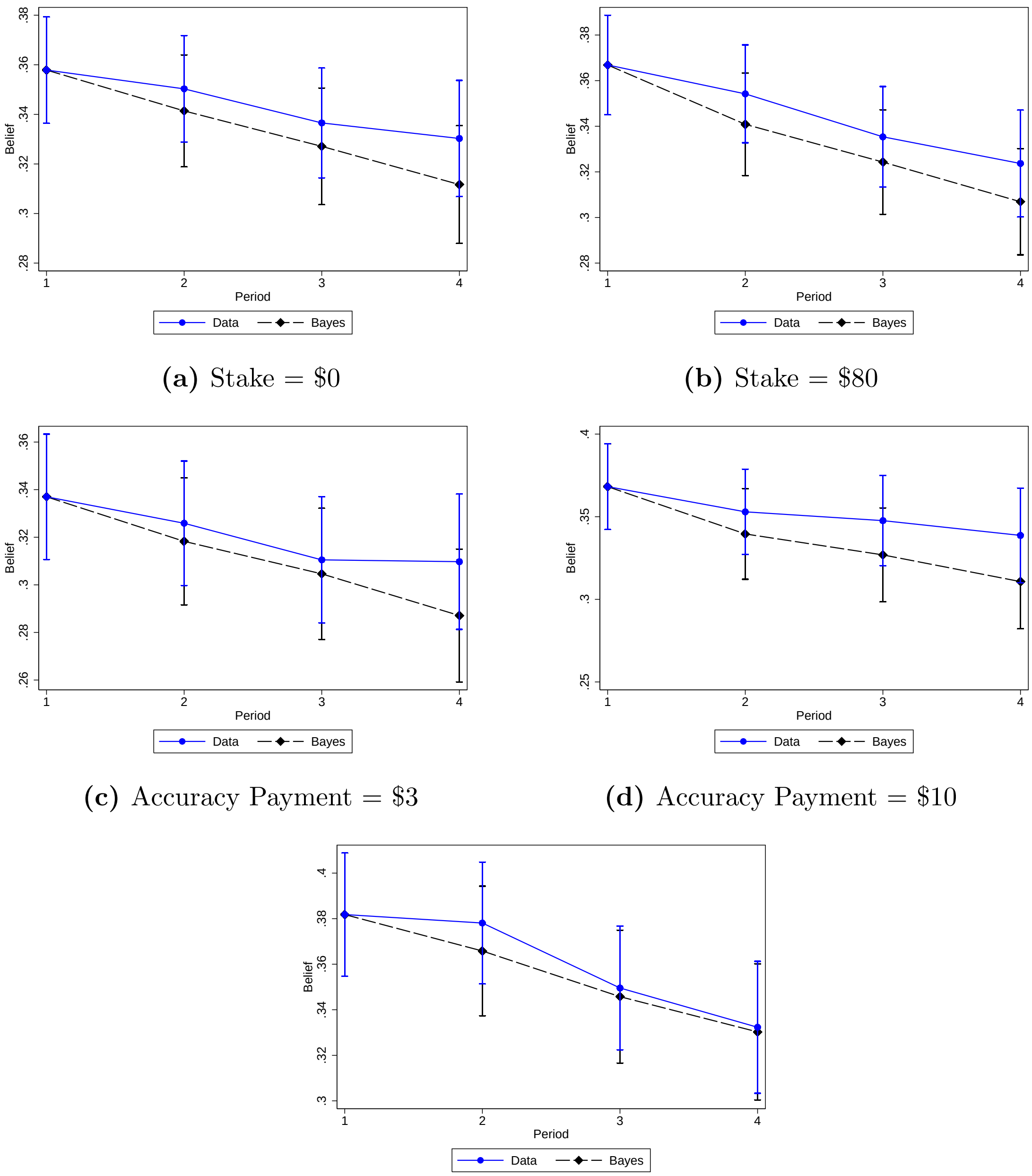
From Figure 7 there does not appear to be any sizeable differences in updating behavior across these different payment conditions. The results on differences between a stake of \$0 versus \$80 are consistent with Barron (2016), who does not find evidence of asymmetry when individuals have a financial stake in an event. Note also that the prior varies slightly by payment conditions; updating patterns by prior are presented in Figure 9 below.

Next, in Figure 8 I present the analogous analysis for each of the four events, with the quiz event split into the self and other treatments. For the two dice events, which involved the probability that particular outcomes from rolls of either two or four dice had occurred, updating appears to be more conservative than the aggregate. The pattern is also seen when individuals estimate the probability that another randomly selected, anonymous individual in the room had scored in the top 15% on the earlier taken quiz (quiz: other performance).

For the quiz (self performance) event, which involved the probability that the individual believed they scored in the top 15% of quiz takers, updating appears to adhere more closely to the Bayesian prediction. This is also true for the weather event, which occurred when subjects had correctly estimated the mean temperature $\pm$ 5 degrees F in New York City on a randomly selected day in the previous calendar year. In the aggregate, unlike the findings of Mobius et al. (2014) and Eil and Rao (2011), updating about own performance does not appear to deviate much from the Bayesian prediction.

Finally, Figure 9 presents similar figures for different values of the prior. Updating appears conservative for low values of the prior, well calibrated for moderate values, and too responsive (exhibiting base rate neglect) for high values of the prior. These patterns are highly suggestive that some of the differences in updating observed across events are driven by differences in average values of the prior, rather than differences in the events themselves. In particular, elicited priors for the two dice events and the quiz (other) event are significantly lower than those for the weather event and the quiz (self) event, and additionally exhibit substantially greater levels of conservatism.

Figure 7: Evolution of Beliefs By Stake and Accuracy Conditions



(a) Stake = \$0

(b) Stake = \$80

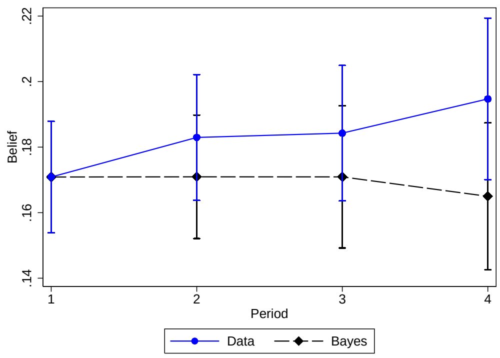
(c) Accuracy Payment = \$3

(d) Accuracy Payment = \$10

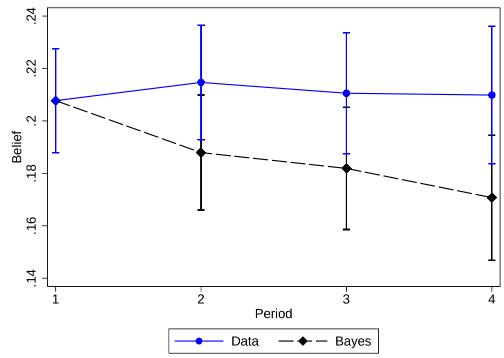
(e) Accuracy Payment = \$20

The path of beliefs starting from the prior (period 1), and after each sequential signal (periods 2 through 4). Average individual responses are the blue solid line, the Bayesian benchmark is marked as the black dashed line. Bayesian benchmark takes prior beliefs, and subsequently uses Bayes' rule to update beliefs. Error bands represent 95% confidence intervals. Note the potential difference in the range of prior beliefs, on the vertical axis. $N = \{646, 634, 424, 436, 420\}$ per round, respectively for (a)-(e).

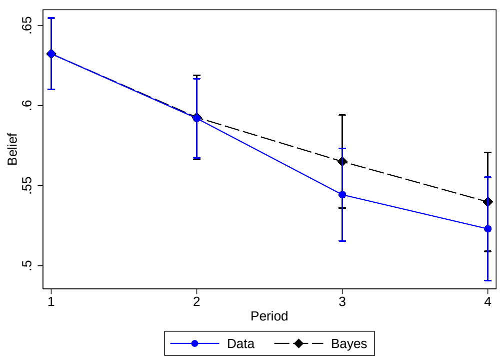
Figure 8: Evolution of Beliefs: Individual Domains



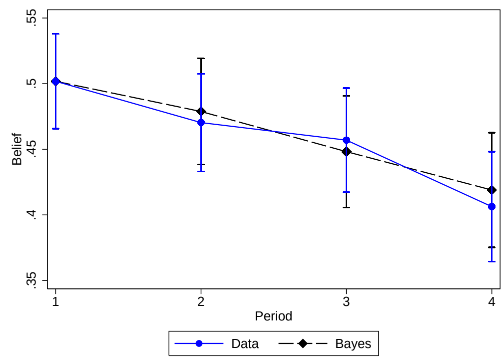
(a) Easy Dice



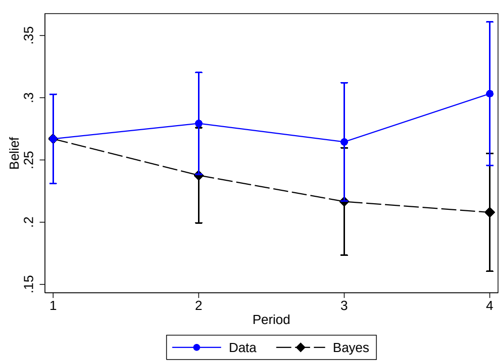
(b) Hard Dice



(c) Weather Event



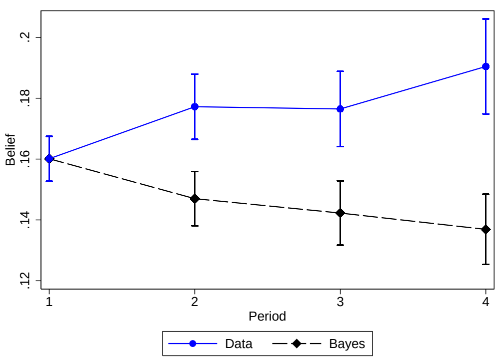
(d) Quiz Event (self performance)



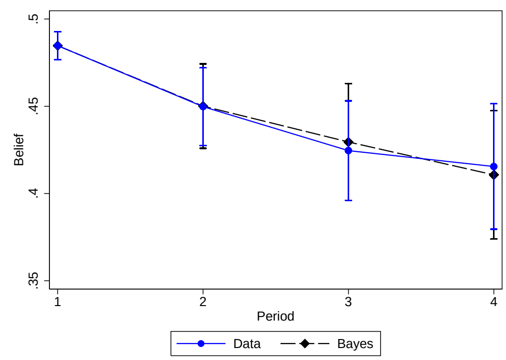
(e) Quiz Event (other performance)

The path of beliefs starting from the prior (period 1), and after each sequential signal (periods 2 through 4). Average individual responses are the blue solid line, the Bayesian benchmark is marked as the black dashed line. Bayesian benchmark takes prior beliefs, and subsequently uses Bayes' rule to update beliefs. Error bands represent 95% confidence intervals. Note the difference in the range of prior beliefs, on the vertical axis. $N = \{318, 318, 326, 223, 95\}$ per round, respectively for (a)-(e).

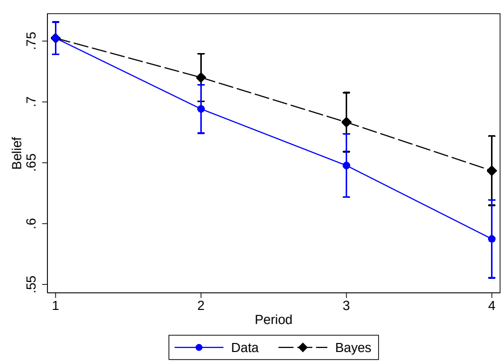
Figure 9: Evolution of Beliefs By Value of the Prior



(a) Prior $\leq$ 40%



(b) 40% $\leq$ Prior $\leq$ 60%



(c) 60% $\leq$ Prior

All events and all stake conditions, sample restricted to individuals updating with indicated prior. The path of beliefs starting from the prior (period 1), and after each sequential signal (periods 2 through 4). Average individual responses are the blue solid line, the Bayesian benchmark is marked as the black dashed line. Error bands represent 95% confidence intervals. $N = \{798, 185, 297\}$ average per round, respectively for (a)-(c).

6.3 Updating Framework: By Event/Stake/Accuracy Payment

Here I replicate the primary analysis found in Table reftab:primary , looking at each of the financial stake and accuracy payment conditions separately. As can be seen in Table 5, there is no clear pattern that emerges within either the accuracy payment or within the financial stake conditions respectively. A formal statistical test confirms that I cannot reject equality between the $0 and $80 financial stake conditions, nor between the $3, $10, and $20 accuracy payment conditions. This analysis suggests that different payments for accuracy do not alter updating behavior. Similarly, holding a large financial stake in an event does not alter updating behavior relative to holding no stake.

Table 5: Updating Beliefs for All Events: By Accuracy Payment and Financial Stake

Regressor	Stake = 0	Stake = 80	Acc = 3	Acc = 10	Acc = 20	Total
δ	0.910***	0.918***	0.920***	0.922***	0.898***	0.914***
	(0.012)	(0.014)	(0.017)	(0.014)	(0.016)	(0.009)
β_1	0.587***	0.588***	0.560***	0.662***	0.540***	0.588***
	(0.045)	(0.043)	(0.054)	(0.063)	(0.059)	(0.034)
β_0	0.807***	0.780***	0.774***	0.749***	0.861*	0.793***
	(0.047)	(0.047)	(0.066)	(0.060)	(0.074)	(0.038)
P-Value ($\delta = 1$)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P-Value ($\beta_1 = 1$)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
P-Value ($\beta_0 = 1$)	0.0001	0.0000	0.0008	0.0001	0.0616	0.0000
P-Value ($\beta_1 = \beta_0$)	0.0001	0.0011	0.0042	0.2112	0.0000	0.0000
R^2	0.83	0.84	0.84	0.84	0.82	0.84
Observations	1704	1656	1128	1143	1089	3360

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

In Table 6 I replicate the primary analysis found in Table 2, but excluding any observations where an individual had a financial stake of $80. There do not appear to be any consistent differences in this subsample. In the final column of Table 6 I use the same sampling procedure as Mobius et al. (2014), in order to provide a more comparable estimation to their study for the Quiz event. One can see that the sampling procedure does not significantly alter the pattern of observed results.

Table 6: Updating Beliefs Within Events: No Financial Stake Only

Regressor	Easy Dice	Hard Dice	Weather	Quiz (S)	Quiz (O)	Quiz (M. et al)
δ	0.839***	0.897***	0.909***	0.924**	0.894*	0.918**
	(0.049)	(0.028)	(0.027)	(0.030)	(0.055)	(0.035)
β_1	0.317***	0.430***	0.683***	0.616***	0.816	0.714***
	(0.146)	(0.092)	(0.089)	(0.078)	(0.200)	(0.090)
β_0	1.073	0.815*	0.783***	0.799**	0.778	0.917
	(0.171)	(0.109)	(0.067)	(0.087)	(0.176)	(0.099)
P-Value ($\delta = 1$)	0.0013	0.0004	0.0009	0.0138	0.0603	0.0226
P-Value ($\beta_1 = 1$)	0.0000	0.0000	0.0005	0.0000	0.3623	0.0021
P-Value ($\beta_0 = 1$)	0.6711	0.0906	0.0014	0.0224	0.2144	0.4049
P-Value ($\beta_1 = \beta_0$)	0.0095	0.0136	0.3791	0.0472	0.9040	0.0521
R^2	0.66	0.77	0.73	0.83	0.79	0.84
Observations	435	421	447	294	107	225

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant. Includes only updated beliefs about events where individuals did not hold any additional financial stake in the outcome.

6.4 Invariance, Sufficiency, and Stability

In this section I investigate three additional properties that are satisfied when updated beliefs follow Bayes' rule. First, the structure of Bayes' rule implies a sufficiency condition, that priors are sufficient statistics for all the information contained in past signals. In other words, after controlling for prior beliefs, lagged information does not significantly predict posterior beliefs. To examine whether updating behavior can be shown to satisfy the sufficiency condition I follow Mobius et al. (2014) and include lagged signals as independent variables. Table 7 shows the regressions that include these lagged signals, using only actively revised beliefs.[33] There is some evidence that overall, the updating process may not satisfy the sufficiency condition, as the first signal received has a significant effect on belief updating in round 3.[34]

The next property Bayes' rule satisfies is stability: that updating remains stable across time. Looking across the three updating rounds in Table 1, there appear to be differences. Overall, I can reject equality across rounds 1 to 3 for δ at the 1% level for the sample of actively revised beliefs, but I cannot reject equality for the response to signals β_1 and β_0. Thus there is some evidence that updating is not stable, primarily with respect to δ, the

[33] The test is misspecified if I include posterior beliefs that are not updated. Nonetheless, the results are not substantively affected conducting the test on the full sample.

[34] While Mobius et al. (2014) do not reject sufficiency, it is worth noting that the ratio of the values of coefficients on lagged signals to current signals is of the same magnitude.

Table 7: Examining Sufficiency

Dependent Variable: Belief		
Regressor	Round 2	Round 3
δ	0.890***	0.880***
	(0.027)	(0.023)
β_1	1.030***	1.247***
	(0.065)	(0.074)
β_0	1.287***	1.347***
	(0.064)	(0.066)
β_{t-1}	0.052	0.048
	(0.045)	(0.042)
β_{t-2}		0.164***
		(0.042)
R^2	0.82	0.82
Observations	640	670

Analysis uses OLS regression. Difference is *significant from zero* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant. The sample is restricted to include only subjects who actively revised their beliefs in the direction predicted by Bayes' rule. β_{t-k} refers to the k^{th} lagged signal.

weight given to the prior.[35]

Finally, the invariance property is said to hold when $\delta = 1$, that is the change in logit beliefs depends only on past signals. $\delta = 1$ is rejected in the data at the 1% level. However, despite these three conditions not being met in the data, it is important to note that the magnitude of this deviation is reasonably small, in the sense that the resulting posteriors are very close to their Bayesian counterparts.[36] This is because the magnitude of these failures is small in relative terms. The coefficients are close to one, and even the largest response to past signals is only 14% of the response to a current signal. The overall result remains that posteriors are close in magnitude to the Bayesian prediction.

[35] If I examine the sample including beliefs that are not actively revised, I can reject equality for each of the three coefficients δ, β_1, β_0 at the 5% level respectively.

[36] $\delta = 0.873$ is reasonably close to 1, indicating that invariance is close to being satisfied in this sample. As in Mobius et al. (2014) a concern is that β_1 and β_0 are functions of prior beliefs, but that these effects cancel out to give a coefficient of δ approximately equal to 1. To examine if this is a potential issue in this data I check whether there are significant interaction effects between receiving affirmative signals, and the prior belief. These interactions are never significant at any reasonable significance level, indicating that this is not a problem for the data.

6.5 Sampling Robustness Checks

Table 8 examines the impact of how restricting the sample alters updating estimates in the main framework. The first column presents the main analysis (Column 3 in Table 1), but includes observations where belief updates go in the opposite direction that Bayes' rule predicts. The second column replaces boundary observations of 0 or 1 with 0.01 or 0.99 respectively. In Table 1 these were dropped. Finally the third column also truncates boundary observations, and includes updates in the wrong direction. The third column thus presents the full data, with no exclusions.

Table 8: Relaxing Sample Restrictions and Full Sample

Regressor	Include Wrong Dir.	Include Boundary	Include All
δ	0.910***	0.914***	0.914***
	(0.010)	(0.010)	(0.011)
β_1	0.506***	0.727***	0.649***
	(0.033)	(0.045)	(0.045)
β_0	0.714***	0.903**	0.805***
	(0.038)	(0.045)	(0.045)
P-Value ($\delta = 1$)	0.0000	0.0000	0.0000
P-Value ($\beta_1 = 1$)	0.0000	0.0000	0.0000
P-Value ($\beta_0 = 1$)	0.0000	0.0306	0.0000
P-Value ($\beta_1 = \beta_0$)	0.0000	0.0003	0.0022
R^2	0.81	0.81	0.79
Observations	3537	3654	3840

Analysis uses OLS regression. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant. First column includes updates in direction predicted by Bayes' rule. Second column replaces boundary probabilities with 0.01 or 0.99 respectively. Third column is entire sample.

6.6 Active Updates Only / Sampling Robustness Checks

Table 9 presents the analysis of Table 1, but restricting the sample to only active updates. The results show that subjects appear to suffer from the opposite bias of conservatism, as they are over-responsive to information. This is largely drive by response to a negative signal, but does not appear to differ between good or bad news, versus just news. As such, symmetry can be rejected at the 1% level.

Table 9: Updating Beliefs for All Events

Active Updates: Reponse to Contemporaneous Signal			
Regressor	Good/Bad News	Just News	All
δ	0.882***	0.863***	0.873***
	(0.018)	(0.022)	(0.014)
β_1	1.060	1.092	1.074
	(0.052)	(0.071)	(0.047)
β_0	1.295***	1.323***	1.305***
	(0.050)	(0.071)	(0.046)
P-Value ($\delta = 1$)	0.0000	0.0000	0.0000
P-Value ($\beta_1 = 1$)	0.2529	0.1962	0.1138
P-Value ($\beta_0 = 1$)	0.0000	0.0000	0.0000
P-Value ($\beta_1 = \beta_0$)	0.0002	0.0121	0.0000
R^2	0.81	0.79	0.80
Observations	1121	799	1920

Analysis uses OLS regression. Includes only active updates, log likelihood calculated for only the current signal. Difference is *significant from 1* at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. R^2 corrected for no-constant.

6.7 Correlates of Active Update Decisions

Table 10: Correlates of Active Updating Decision

	Dependent Variable: Active Update			
Regressor				
Prior	1.596***	1.629***	1.576***	1.642***
	(0.136)	(0.124)	(0.153)	(0.139)
Prior^2	−1.678***	−1.595***	−1.695***	−1.630***
	(0.145)	(0.133)	(0.154)	(0.141)
Event = Hard Dice			−0.012	−0.017
			(0.022)	(0.023)
Event = Weather			0.068**	0.011
			(0.030)	(0.028)
Event = Quiz (self)			0.053*	0.033
			(0.030)	(0.028)
Event = Quiz (other)			0.100***	0.008
			(0.037)	(0.034)
Round 3			0.045***	0.047***
			(0.016)	(0.017)
Round 4			0.137***	0.138***
			(0.019)	(0.019)
Male			0.164**	
			(0.066)	
Percentile Score			0.198***	
			(0.069)	
Male × Percentile Score			−0.233**	
			(0.112)	
Econ Major			−0.014	
			(0.041)	
Constant	0.347***		0.155***	
	(0.025)		(0.048)	
Individual Fixed Effects	NO	YES	NO	YES
R^2	0.06	0.37	0.10	0.38
Observations	3654	3654	3483	3654

Analysis uses OLS regression, dependent variable is binary for whether subject updated prior. Difference is significant from zero at * 0.1; ** 0.05; *** 0.01. Robust standard errors clustered at individual level. Omitted event is Easy Dice. Excludes belief revisions in the opposite direction predicted by Bayes' rule.

References

Ambuehl, Sandro and Shengwu Li, "Belief Updating and the Demand for Information," *mimeo*, 2015.

Antoniou, Constantinos, Glenn W. Harrison, Morten I. Lau, and Daniel Read, "Subjective Bayesian beliefs," *Journal of Risk and Uncertainty*, feb 2015, *50* (1), 35–54.

Armantier, Olivier and Nicolas Treich, "Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging," *European Economic Review*, aug 2013, *62*, 17–40.

Barber, B M and T Odean, "Boys will be boys: Gender, overconfidence, and common stock investment," *Quarterly Journal of Economics*, 2001, *116* (1), 261–292.

Barron, Kai, "Belief updating: Does the 'good-news, bad-news' asymmetry extend to purely financial domains?," *WZB Discussion Paper*, 2016, (October).

Benabou, R. and J. Tirole, "Self-Confidence and Personal Motivation," *The Quarterly Journal of Economics*, aug 2002, *117* (3), 871–915.

Blanco, Mariana, Dirk Engelmann, Alexander K. Koch, and Hans-Theo Normann, "Belief elicitation in experiments: is there a hedging problem?," *Experimental Economics*, jul 2010, *13* (4), 412–438.

Buser, Thomas, Leonie Gerhards, and Joël J Van Der Weele, "Measuring Responsiveness to Feedback as a Personal Trait," *Tinbergen Institute Discussion Paper*, 2016.

Camerer, Colin F, "Do Biases in Probability Judgment Matter in Markets? Experimental Evidence," *The American Economic Review*, 1987, *77* (5), 981–997.

_ , "Individual Decision Making," in John H. Kagel and Alvin E. Roth, eds., *The Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press, 1995, pp. 587–703.

Clark, Jeremy and Lana Friesen, "Overconfidence in forecasts of own performance: An experimental study," *Economic Journal*, 2009, *119* (534), 229–251.

Coutts, Alexander, "Testing Models of Belief Bias: An Experiment," *mimeo*, 2015.

Eil, David and Justin M Rao, "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics*, may 2011, *3* (2), 114–138.

Ertac, Seda, "Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback," *Journal of Economic Behavior & Organization*, dec 2011, *80* (3), 532–545.

Gilboa, Itzhak and David Schmeidler, *A Theory of Case-Based Decisions*, Cambridge: Cambridge University Press, 2001.

Grether, David M., "Bayes Rule as a Descriptive Model: The Representativeness Heuristic," *The Quarterly Journal of Economics*, nov 1980, *95* (3), 537.

— , "Testing bayes rule and the representativeness heuristic: Some experimental evidence," *Journal of Economic Behavior & Organization*, jan 1992, *17* (1), 31–57.

Grossman, Zachary and David Owens, "An unlucky feeling: Overconfidence and noisy feedback," *Journal of Economic Behavior & Organization*, nov 2012, *84* (2), 510–524.

Healy, Paul J., "Explaining the BDM - Or Any Random Binary Choice Elicitation Mechanism - To Subjects," *mimeo*, 2016.

Holt, Charles and Angela M. Smith, "An update on Bayesian updating," *Journal of Economic Behavior & Organization*, feb 2009, *69* (2), 125–134.

Kahneman, Daniel and Amos Tversky, "On the psychology of prediction.," *Psychological Review*, 1973, *80* (4), 237–251.

Karni, Edi, "A Mechanism for Eliciting Probabilities," *Econometrica*, 2009, *77* (2), 603–606.

— **and Zvi Safra**, "The impossibility of experimental elicitation of subjective probabilities," *Theory and Decision*, may 1995, *38* (3), 313–320.

— , — , and The Econometric Society, "Preference Reversal" and the Observability of Preferences by Experimental Methods," *Econometrica*, 1987, *55* (3), pp. 675–685.

Kuhnen, Camelia M, "Asymmetric Learning from Financial Information," *The Journal of Finance*, 2014, *LXX* (5), 2029–2062.

Landier, Augustin, "Wishful thinking a model of optimal reality denial," 2000.

Machina, Mark J. and David Schmeidler, "A More Robust Definition of Subjective Probability," *Econometrica*, jul 1992, *60* (4), 745.

Mobius, Markus, Muriel Niederle, Tanya Rosenblat, and Paul Niehaus, "Managing Self-Confidence : Theory and Experimental Evidence," *mimeo*, 2014.

Scheier, Michael E and Charles S. Carver, "Dispositional Optimism and Physical Well-Being: The Influence of Generalized Outcome Expectancies on Health," *Journal of Personality*, jun 1987, *55* (2), 169–210.

Schlag, Karl H., James Tremewan, and Joël J. van der Weele, "A penny for your thoughts: a survey of methods for eliciting beliefs," *Experimental Economics*, 2014.

Schotter, Andrew and Isabel Trevino, "Belief Elicitation in the Laboratory," *Annual Review of Economics*, aug 2014, *6* (1).

Shah, Punit, Adam J L Harris, Geoffrey Bird, Caroline Catmur, and Ulrike Hahn, "A pessimistic view of optimistic belief updating," *Cognitive Psychology*, 2016, *90*, 71–127.

Sharot, T, Cw Korn, and Rj Dolan, "How unrealistic optimism is maintained in the face of reality," *Nature Neuroscience*, 2011, *14* (11), 1475–1479.

Sharot, T., R. Kanai, D. Marston, C. W. Korn, G. Rees, and R. J. Dolan, "Selectively altering belief formation in the human brain," *Proceedings of the National Academy of Sciences*, 2012, *109* (42), 17058–17062.

Wiswall, Matthew and Basit Zafar, "How Do College Students Respond to Public Information about Earnings?," *Journal of Human Capital*, jun 2015, *9* (2), 117–169.

Yariv, Leeat, "I'll See It When I Believe It - A Simple Model of Cognitive Consistency," *mimeo*, 2005.