

Belief Updating: Does the ‘Good-news, Bad-news’ Asymmetry Extend to Purely Financial Domains?*

Kai Barron[†]

July 15, 2019

Abstract

Bayes’ statistical rule remains the status quo for modeling belief updating in both normative and descriptive models of behavior under uncertainty. Some recent research has questioned the use of Bayes’ rule in descriptive models of behavior, presenting evidence that people overweight ‘good news’ relative to ‘bad news’ when updating ego-relevant beliefs. In this paper, we present experimental evidence testing whether this ‘good-news, bad-news’ effect is present in a financial decision making context (i.e. a domain that is important for understanding much economic decision making). We find no evidence of asymmetric updating in this domain. In contrast, in our experiment, belief updating is close to the Bayesian benchmark on average. However, we show that this average behavior masks substantial heterogeneity in individual updating behavior. We find no evidence in support of a sizeable subgroup of asymmetric updaters.

JEL Classification: C11, C91, D83

Keywords: economic experiments, Bayes’ rule, belief updating, belief measurement, proper scoring rules, motivated beliefs.

*I would like to thank Syngjoo Choi, Alex Coutts, Steffen Huck, Terri Kneeland, Peter Schwardmann, Yves Le Yaouanq, Lukas Wenner, and the audiences at RES, ECBE, THEEM, ASFEE, ENTER Jamboree, and the ESA conferences for comments and helpful discussions, as well as Brian Wallace, Lukas Wenner, Roberta De Filippis, and Nina Bonge for their support in running the experiments. Kai Barron gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft (through CRC TRR 190). All mistakes are my own.

[†]WZB, Reichpietschufer 50, 10785 Berlin, Germany; e-mail: kai.barron@wzb.eu

1 Introduction

Throughout our lives, we are constantly receiving new information about ourselves and our environment. The way that we filter, summarize and store this new information is of critical importance for the quality of our decision making. Theories of human behavior under dynamic uncertainty are therefore enriched by an understanding of how individuals process new information. Economists typically write down models where information is summarized in the form of probabilistic ‘beliefs’ over states of the world, and updated upon receipt of new information according to Bayes’ rule. However, the assumption that individuals process information in a statistically accurate way is receiving an increasing amount of attention, with many studies documenting systematic deviations.¹ One important strand of this literature examines whether belief formation and updating is influenced by the *affective content* of the new information², i.e. whether individuals update their beliefs symmetrically in response to ‘good-news’ and ‘bad-news’ (see, for example, [Eil and Rao \(2011\)](#); [Ertac \(2011\)](#); [Möbius et al. \(2014\)](#); [Coutts \(2019\)](#)).³

Essentially, this literature tests an implicit assumption of Bayesian updating, namely that the only object that is relevant for predicting an individual’s belief is her information set, and therefore her beliefs are completely unaffected by the prizes and punishments she will receive in different states of the world. This fundamental assumption - that people update their beliefs symmetrically - is of paramount importance because it underpins the results of a wide range of theoretical studies within economics, including all research in which agents receive new information and form rational expectations. As noted by [Brunnermeier and Parker \(2005\)](#), since [Muth \(1960, 1961\)](#) and [Lucas \(1976\)](#), the vast majority of economic research involving uncertainty has built on the rational expectations assumption, with diverse applications including the vast literatures on *social learning* (see, for example, [Chamley \(2004\)](#) for a review), *capital markets* (e.g. [Fama \(1970, 1976\)](#)), *intertemporal portfolio choice problems* (e.g. [Mossin \(1968\)](#); [Merton \(1969\)](#); [Samuelson \(1969\)](#); [Lewellen and Shanken \(2002\)](#)) and *consumption savings problems* (e.g. [Friedman \(1957\)](#); [Hall \(1978\)](#)). Therefore, it is essential to empirically test whether the assumption that people form beliefs symmetrically in response to ‘good-news’ and ‘bad-news’, such that their posterior

¹Some notable examples include the *representativeness bias* ([Grether, 1978, 1980, 1992](#)), *cognitive dissonance* ([Akerlof and Dickens, 1982](#)), *anticipatory utility* ([Loewenstein, 1987](#); [Caplin and Leahy, 2001](#); [Brunnermeier and Parker, 2005](#)), *base rate neglect* ([Kahneman and Tversky, 1973](#); [Holt and Smith, 2009](#)), *confirmatory bias* ([Rabin and Schrag, 1999](#)), *motivated belief formation* ([Benabou and Tirole, 2002](#)), *correlation neglect* ([Enke and Zimmermann, 2019](#)), and *selection neglect* ([Esponda and Vespa, 2018](#); [Barron et al., 2019](#); [Enke, 2019](#)).

²As an indication of the importance of this strand of the literature, in their recent paper taking stock of the current state of the motivated reasoning literature, [Bénabou and Tirole \(2016\)](#) point to *asymmetric updating* as one of the main testable implications of motivated reasoning.

³This literature takes inspiration from the fact that, in many circumstances, individuals seem to display optimistic or overconfident beliefs (see, e.g., [Malmendier and Taylor \(2015\)](#) and other papers in the “Symposium on Overconfidence” in the Fall 2015 issue of the *Journal of Economic Perspectives*). This leads naturally to the question of how these beliefs are sustained when receiving a constant stream of feedback. One potential explanation for optimistically biased beliefs persisting in the face of feedback is that belief updating operates asymmetrically in response to ‘good-news’ and ‘bad-news’.

beliefs are completely unaffected by their hopes and wishes, is consistent with how belief updating actually works.

With this objective in mind, in this paper we employ a laboratory experiment to study how individuals update their beliefs from exogenously assigned prior beliefs in a two-state world when they receive a sequence of partially informative binary signals. In particular, we vary the financial rewards associated with one of the two states of the world in order to test whether these state-contingent financial rewards influence belief updating. A nice feature of this experimental design is that it allows us to compare posterior beliefs in situations where the entire information set is held constant, but the rewards associated with the states of the world are varied. For example, we can compare how two groups of individuals revise their beliefs when both groups share the same prior belief and receive an equally informative signal, but for one group of individuals the signal is ‘good news’ and for others the signal constitutes ‘bad news’. Furthermore, we can conduct a similar exercise for a single individual, by comparing two situations in which she has the same information set, and receives the same signal, but the signal constitutes ‘good news’ in one instance, and ‘bad news’ in the other. Our experimental design therefore permits a clean test of the *asymmetric updating hypothesis*.

In our experiment we study belief updating in two contexts. In our SYMMETRIC treatment, we examine how subjects update their beliefs when they have an equal stake in each of the underlying states and are therefore indifferent about which state is realized. We compare this with updating behavior in our two ASYMMETRIC treatments, in which a larger bonus payment will be paid if one state of the world is realized. Here, one would expect that subjects prefer that the state with the bonus payment is realized. In all treatments, subjects are told the true prior probability of the two states, and then update their beliefs upon receipt of a sequence of informative but noisy signals.⁴ We elicit this sequence of beliefs. In our experiment, we can therefore conduct two separate tests of the asymmetric updating hypothesis. Firstly, we can compare how the same individual responds to ‘good-news’ and ‘bad-news’ *within* the ASYMMETRIC treatments. Secondly, we can conduct a *between-subject* comparison of belief updating in the SYMMETRIC treatment and ASYMMETRIC treatments. Each individual in our experiment faces only one incentive environment. However, since we exogenously endow participants with a prior over the states of the world, we are able to repeat the exercise several times for each individual and study how they update from each of five different priors, p_0 , chosen from the set $\{\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}\}$.

The experimental design and analysis aim to address several challenges that are present when studying belief updating in the presence of state dependent stakes. Firstly, we use exogenous variation in the priors to ensure that the estimates are robust to the econometric issues that arise when a right-hand side variable (i.e. the prior) is a lagged version of the dependent variable (i.e.

⁴The ASYMMETRIC treatments correspond to the setup studied in Möbius et al. (2014) and Eil and Rao (2011), where a ‘good’ underlying state is being higher in the IQ or attractiveness distribution in relation to the other participants in the experiment.

the posterior). Secondly, we avoid a second type of endogeneity issue, which arises when the underlying states are defined as a function of some personal characteristic of the individual (e.g. her relative IQ) that might also be related to how she updates (see Appendix C for further details). Thirdly, we measure the influence that hedging has on belief elicitation when there are state-dependent stakes. Furthermore, we conduct several exercises to correct our estimates for this hedging influence - both experimentally, and econometrically. Fourth, our experimental design allows us to study belief updating from priors spanning much of the unit interval. Importantly, averaging across all subjects, the design generates a balanced distribution of ‘good’ and ‘bad’ signals.

The empirical strategy employed permits testing for several commonly hypothesised deviations from Bayesian updating, including confirmatory bias and base rate neglect, however the focus of both the experimental design, and the analysis, is on testing for the presence of an asymmetry in updating. Our results show no evidence in favour of asymmetric updating in response to ‘good-news’ in comparison to ‘bad-news’ in the domain of financial outcomes (when there is a minimal role for ego-utility). Several robustness exercises are carried out in support of this conclusion. Furthermore, we find that average updating behavior is well approximated by Bayes’ rule.⁵ This average behavior masks substantial heterogeneity in updating behavior at the individual level, but we find no evidence in support of a sizeable subgroup of asymmetric updaters.

The evidence reported here contributes to the recent literature studying the asymmetric updating hypothesis across different contexts ([Eil and Rao \(2011\)](#), [Ertac \(2011\)](#), [Grossman and Owens \(2012\)](#), [Mayraz \(2013\)](#), [Möbius et al. \(2014\)](#), [Kuhnen \(2015\)](#), [Schwardmann and Van der Weele \(2016\)](#), [Gotthard-Real \(2017\)](#), [Charness and Dave \(2017\)](#), [Heger and Papageorge \(2018\)](#), [Coutts \(2019\)](#) and [Buser et al. \(2018\)](#)). The results in this literature thus far are surprisingly heterogeneous, with some papers finding an greater responsiveness to good-news; some a greater responsiveness to bad-news; and some no evidence of an asymmetry.⁶ There are several candidate contextual and experimental design factors that could potentially be driving these heterogeneous results. Section 7 offers a detailed discussion of these factors, and asks whether the existing body of evidence can help us to detect a systematic pattern that organises the results (for an alternative discussion of this literature, see [Benjamin \(2019\)](#), who taxonomizes the existing evidence on errors in probabilistic reasoning).

The remainder of the paper proceeds as follows. Section 2 outlines the theoretical framework, Section 3 details the experimental design, Section 4 provides some descriptive statistics, Section 5 presents the empirical specification, Section 6 discusses the related literature and Section 7 concludes.

⁵This is in line with [Holt and Smith \(2009\)](#) and [Coutts \(2019\)](#), who find that average posteriors across all individuals are well approximated by posteriors obtained by applying Bayes’ rule.

⁶See the Discussion section below for further details.

2 Theoretical Framework and Hypotheses

In the following section, we discuss a simple framework for belief updating that augments the standard benchmark of Bayesian updating by allowing for several of the deviations from Bayes' rule commonly discussed in the psychology and related economics literature. This framework is borrowed from Möbius et al. (2014) and is currently a common approach used in this type of descriptive belief updating study. The basic idea is that, while Bayes' rule captures the normative benchmark for how we might think a rational agent *should* update her beliefs⁷, it has been argued that, descriptively, people may update their beliefs in ways that depart systematically from Bayes' rule.⁸ The framework below serves to facilitate a discussion of these different potential deviations, and motivates the empirical approach that we will use to test whether they are observed in our data.

In short, below we describe a model that embeds the normative Bayesian benchmark, but also allows for commonly discussed belief updating distortions. The aim will be to then make use of this model to test whether subjects update their beliefs like a Bayesian automaton or display some systematic deviations from this Bayesian benchmark. Most importantly, the model permits a discussion of what observed updating behavior should look like if agents update their beliefs asymmetrically in response to 'good-news' and 'bad-news'.

2.1 A Simple Model of Belief Formation

We consider a single agent who forms a belief over two states of the world, $\omega \in \{A, B\}$, at each point in time, t . One of these states of the world is selected by nature as the 'correct' (or 'realized') state, where state $\omega = A$ is chosen with prior probability \bar{p}_0 (known to the agent). The agent's belief at time t is denoted by $\pi_t \in [0, 1]$, where π_t is the agent's belief regarding the likelihood that the state is $\omega = A$ and $1 - \pi_t$ is the agent's belief that the state is $\omega = B$. In each period, the agent receives a signal, $s_t \in \{a, b\}$, regarding the state of the world, which is correct with probability $q \in (\frac{1}{2}, 1)$. In other words, $p(a|A) = p(b|B) = q > \frac{1}{2}$. Furthermore, the history, H_t , is defined

⁷This statement is not uncontentious. One argument in favor of using Bayes' rule as a normative benchmark for belief updating is that Bayes' rule captures the objective statistical relationship between a prior probability and a posterior probability, given new information. Therefore, if an individual updates her beliefs according to Bayes' rule, she will always hold beliefs that are as accurate as possible, given her information set. This will allow her to use these optimal beliefs to guide her decision making. However, this argument that Bayes' rule represents the normative optimum rests on the assumption that beliefs serve only an *instrumental* role in guiding decision making. If, for example, we relax this assumption and allow beliefs to yield *intrinsic* utility, then this argument no longer holds as it may be optimal for the decision maker to hold beliefs that are distorted away from the Bayesian posterior.

⁸One argument used to justify the use of Bayes' rule in descriptive models of human behavior is that evolution should have selected individuals who were able to process new information in a statistically efficient way over individuals who could not. Therefore, Bayesian individuals would have had more accurate beliefs guiding their decision making, and been more likely to survive.

as the sequence of signals received by the agent in periods $1, \dots, t$, with $H_0 = \emptyset$. Therefore, the history at time t is given by $H_t = (s_1, \dots, s_t)$.

In order to study how individuals update their beliefs, we follow Möbius et al. (2014) in considering the following model of *augmented Bayesian updating*:⁹

$$\text{logit}(\pi_{t+1}) = \delta \text{logit}(\pi_t) + \gamma_a \log\left(\frac{q}{1-q}\right) \cdot 1(s_{t+1} = a) - \gamma_b \log\left(\frac{q}{1-q}\right) \cdot 1(s_{t+1} = b) \quad (1)$$

The parameters δ , γ_a and γ_b serve, firstly, to provide structure for a discussion of different ways in which our agent may depart from Bayesian updating, and secondly, to provide a clear prescription for how to test for these departures in our empirical analysis below.

If $\delta = \gamma_a = \gamma_b = 1$ then the agent updates her beliefs according to Bayes' rule. However, if we consider deviations from this benchmark, we see that δ captures the degree to which the magnitude of the agent's prior affects her updating. For example, if $\delta > 1$ then this suggests that the agent displays a *confirmatory bias*¹⁰, whereby she is more responsive to information that supports her prior. In contrast, $\delta < 1$ suggests she is more responsive to information that contradicts her prior (i.e. *base rate neglect*¹¹). The former would predict that beliefs will polarize over time, while the latter would predict that over time beliefs remain closer to 0.5 in a two-state world than Bayes' rule would predict.

The parameters, γ_a and γ_b capture the agent's responsiveness to information. If $\gamma_a = \gamma_b < 1$ then the agent is less responsive to the information that she receives than a Bayesian updater would be. And if $\gamma_a = \gamma_b > 1$, then she is more responsive than a Bayesian. For example, if $\gamma_a = 2$, then whenever the agent receives a signal $s_t = a$, she updates her belief exactly as much as a Bayesian would if he received two a signals, $s_t = \{a, a\}$. The interpretation of the parameters is summarized in the first five rows of Table 1 below.

⁹For a discussion of core properties underlying Bayes' rule, see Appendix B.1. This discussion serves as a motivation for the approach taken in augmenting Bayes' rule in equation 1. For further details regarding the derivation of equation 1, refer to an earlier working paper version of this paper (Barron, 2018), or refer to the comprehensive survey of this topic provided by Benjamin (2019).

¹⁰For a detailed discussion of the *confirmatory bias*, see Rabin and Schrag (1999). Essentially, it is the tendency to weight information that supports one's priors more heavily than information that opposes one's priors. In this case, when one's prior regarding state $\omega = A$ is greater than 0.5, i.e. $\pi_t > 0.5$, a participant who is prone to the *confirmatory bias* weights signals that support state $\omega = A$ more heavily than signals that support state $\omega = B$; and *vice versa* when her prior suggests state $\omega = B$ is more likely, i.e. $\pi_t < 0.5$.

¹¹One can think of *base rate neglect* in this context as the agent forming her beliefs as if she attenuates the influence of her prior belief when calculating her posterior - i.e. acting as if her prior was closer to 0.5 than it actually was.

Table 1: Interpretation of Parameters: A Summary

Belief Updating Distortion	Parameter Values
Bayesian Updating	$\delta = 1, \gamma_a = 1 \text{ and } \gamma_b = 1$
Confirmatory Bias	$\delta > 1$
Base Rate Neglect	$\delta < 1$
Conservatism	$\gamma_j < 1 \text{ for } \forall j \in \{a, b\}$
Overresponsiveness	$\gamma_j > 1 \text{ for } \forall j \in \{a, b\}$
Optimistic updating (in ASYMMETRIC)	$\gamma_a > \gamma_b$
Pessimistic updating (in ASYMMETRIC)	$\gamma_a < \gamma_b$

Affective States

The discussion in the preceding section has so far focused on purely cognitive deviations from Bayes’ rule. The *affect* or desirability of different states of the world has played no role. However, in most situations in which individuals form beliefs, there are some states that yield an outcome that is preferred to the outcomes associated with other states - i.e. there are *good* and *bad* states of the world. For example, an individual would generally prefer to be more intelligent rather than less intelligent, to win a lottery rather than lose, and for the price of assets in her possession to increase rather than decrease. This implies that new information about the state of the world is also often either *good-news* or *bad-news*.

In order to allow for the possibility that individuals update their beliefs differently in response to *good-news* in comparison to *bad-news*, we relax the assumption that belief updating is orthogonal to the *affect* of the information.¹² To do this, assume that each of the two states of the world is associated with a certain outcome—i.e. in state $\omega = A$, the agent receives outcome x_A , and in state $\omega = B$, she receives x_B . There are now two belief updating scenarios:

- Scenario 1 (SYMMETRIC): the agent is indifferent between outcomes (i.e. $x_A \sim x_B$); and
- Scenario 2 (ASYMMETRIC): the agent strictly prefers one of the two outcomes (i.e. $x_A \succ x_B$).

Now, the question of interest is whether the agent will update her beliefs differently in the SYMMETRIC and ASYMMETRIC scenarios. Under the assumption that the agent’s behavior is consistent with the model described above in Equation 1, this involves asking whether the parameters δ, γ_a and γ_b , differ between the two contexts.

To guide our discussion, we consider the following two benchmarks. The first natural benchmark is Bayes’ rule, which prescribes that all three parameters equal 1 in both the SYMMETRIC and

¹²To avoid ambiguity, in the discussion below, the term ‘preference’ is usually used to refer to preferences over sure outcomes—never to a preference ordering over lotteries. We will also sometimes refer to ‘preferring’ one state of the world to another. This simply captures the idea that an individual prefers the realisation of a state in which a *good* outcome is realised.

ASYMMETRIC contexts—statistically efficient updating of probabilities is unaffected by the state-dependent rewards and punishments. According to Bayes’ rule, news is news, independent of its affective content.

Hypothesis 1 (*Bayesian Updating*) *Individuals update their beliefs according to Bayes’ rule. Therefore, $\delta = 1, \gamma_a = 1$ and $\gamma_b = 1$ in both SYMMETRIC and ASYMMETRIC scenarios.*

The second benchmark that we consider is the *asymmetric updating hypothesis*—that individuals respond more to ‘good-news’ than ‘bad-news’. Here, in our simple framework there are two ways to identify *asymmetric updating*.

Firstly, if we only consider the behavior of individuals **within** the ASYMMETRIC scenario, we can ask whether there is an *asymmetry in updating* after signals that favor the more desirable state $\omega = A$ (‘good-news’), relative to signals that favor the less desirable state $\omega = B$ (‘bad-news’). For example, if $\gamma_a > \gamma_b$, this would indicate that the agent updates more in response to ‘good-news’. We refer to such an agent as an *optimistic updater*. Conversely, if we have $\gamma_a < \gamma_b$ then the agent updates more in response to ‘bad-news’. We refer to such an agent as a *pessimistic updater*.¹³

Secondly, if we compare behavior **between** the SYMMETRIC and ASYMMETRIC scenarios, we can ask whether the parameters of equation 1 differ according to the scenario. We use the postscript $c \in \{A, S\}$ to distinguish the parameters in the two scenarios—i.e. δ^S, γ_a^S and γ_b^S in SYMMETRIC and δ^A, γ_a^A and γ_b^A in ASYMMETRIC. In the SYMMETRIC treatment, where the agent is completely indifferent between the two states, there is no reason to expect her updating to be asymmetric. Therefore, we assume that $\gamma_a^S = \gamma_b^S = \gamma^S$. Thus, the difference $\gamma_a^A - \gamma_a^S$ reflects a measure of the increase in the agent’s responsiveness when information is desirable, relative to when information is neutral in terms of its affect. Similarly, $\gamma_b^A - \gamma_b^S$ is a measure of the increase in the agent’s responsiveness when information is undesirable, relative to the case in which information is neutral in affect.

Hypothesis 2 (*Asymmetric Updating*) *Individuals update their beliefs asymmetrically, responding more to good than bad news. Therefore, **within** the ASYMMETRIC scenario, we will observe $\gamma_a^A > \gamma_b^A$. And in a comparison **between** the SYMMETRIC and ASYMMETRIC scenarios, we will observe $\gamma_a^A - \gamma_a^S > 0$ and $\gamma_b^A - \gamma_b^S < 0$. Together, we can summarize the asymmetric updating hypothesis parameter predictions as follows: $\gamma_a^A > \gamma^S > \gamma_b^A$.*

¹³Notice that these definitions of optimistic and pessimistic updating relate only to the asymmetry in responsiveness to signals supporting different states, but depend on how responsive the agent is to the signals she receives relative to Bayes’ rule. Therefore, under the definition given, we can have an optimistic updater who is less responsive to signals supporting the desirable state than a Bayesian (i.e. $1 > \gamma_a > \gamma_b$). This agent is less responsive than a Bayesian to both desirable and undesirable information, but more responsive to desirable information relative to undesirable information.

In our experiment, we will provide evidence on both of these questions. Firstly, we will examine whether there is an asymmetry in updating behavior within the `ASYMMETRIC` context; and secondly, we will examine whether the parameters of the updating process differ between the `SYMMETRIC` and the `ASYMMETRIC` contexts, in a between subjects comparison. Furthermore, our experiment will allow us to test for other systematic deviations from Bayes’ rule, such as those mentioned in the discussion above. Table 1 summarizes the interpretations of the different values that the belief updating parameters may take.

2.2 Belief Elicitation and Incentives

In order to empirically test the hypotheses above using an experiment, we would like to be able to elicit our participants’ true beliefs. One common feature of many studies using belief elicitation is to assume that participants evaluate prospects using a specific model, and then to show that under the incentives of the elicitation technique, a participant who is following the assumed model should report her true belief (i.e. under the model, truthful revelation is *incentive compatible*).

However, in the context of studying the relationship between preferences and beliefs, two concerns may be raised regarding this approach: firstly, the assumption that participants follow expected utility involves a stronger assumption than we might wish to make; and secondly, the inherent hedging motive faced by participants who have a stake in one state of the world poses an additional challenge (see [Karni and Safra \(1995\)](#) for a discussion). Therefore we adopt the approach developed by [Offerman et al. \(2009\)](#), and extended to accommodate state-dependent stakes as in [Kothiyal et al. \(2011\)](#).

The central idea behind this approach is to acknowledge that the incentive environment within which we elicit beliefs in the laboratory may exert a distortionary influence on the beliefs which some participants report, relative to the beliefs they actually hold, and then measure this distortionary influence of the incentive environment in a separate part of the experiment. Once we have constructed a mapping from true beliefs to reported beliefs within the relevant incentive environment, we can use this function to recover the participant’s true beliefs from her reported beliefs. In other words, our objective is to recover the function that each individual uses to map her true beliefs to the beliefs that she reports within the given incentive environment.

The incentive environment that we will use in our experiment to elicit beliefs is the quadratic scoring rule (QSR).¹⁴ In Appendix B.2, we provide a detailed discussion of way in which reported

¹⁴There are several reasons for adopting this approach: firstly, the QSR has the advantage that it ensures that the decision environment is clear and simple for the participants - essentially they are making a single choice from a list of binary prospects; secondly, the quadratic scoring rule has been commonly used in the literature, with both the *theoretical properties* and *empirical performance* having been studied in detail (see, e.g., [Armantier and Treich \(2013\)](#)); thirdly, in a horse race between elicitation methods, [Trautmann and van de Kuilen \(2015\)](#) show that there is no improvement in the empirical performance of more complex elicitation methods over the [Offerman et al.](#)

beliefs might be distorted under the quadratic scoring rule, depending on the assumed underlying model of preferences. A key example is that it is well documented under the QSR a risk averse agent should distort their reported belief towards 50%. However, this is only true in the absence of state-dependent stakes. In our experiment, we will also be interested in eliciting beliefs when participants have an exogenous stake associated with one of the two states.

In the context of state-dependant stakes, a risk averse EU maximizer will face two distortionary motives in reporting her belief: (i) she will face the motive to distort her belief towards 50% as discussed above; and (ii) in addition, there is a hedging motive, which will compel a risk averse individual to *lower* her reported belief, r_t , towards 0% as the size of the exogenous stake increases.¹⁵

If the participants in our experiment are *risk neutral expected utility* maximizers, the reported beliefs, r_t , that we elicit under the QSR will coincide with their true beliefs, π_t . However, in order to allow for choice behavior consistent with a wider range of decision models, we measure the size of the distortionary influence of the elicitation incentives at an individual level and correct the beliefs accordingly.

A Non-EU ‘Truth Serum’

The [Offerman et al. \(2009\)](#) approach proposes correcting the reported beliefs for the risk aversion caused by the curvature of the utility function or by non-linear probability weighting. This approach involves eliciting participants’ reported belief parameter, r , for a set of risky events where they know the objective probability, p (*known probability*). This is done under precisely the same QSR incentive environment in which the participants’ subjective beliefs, π , regarding the events of interest (where they don’t know the objective probability: *unknown probability*) are elicited. If a subject’s reported beliefs, r , differ from the known objective probabilities, p , this indicates that the subject is distorting her beliefs due to the incentive environment (e.g. due to risk aversion). The objective of the correction mechanism is therefore to construct a map, R , from the objective beliefs, $p \in [0, 1]$, to the reported beliefs, r , for each individual under the relevant incentive environment.

In Appendix B.2, we offer a detailed discussion of how the [Offerman et al. \(2009\)](#) method operates, describe the underlying assumptions, and demonstrate how it can be augmented (as pro-

[\(2009\)](#) method, neither in terms of internal validity, nor in terms of behavior prediction. Out of the set of alternative elicitation techniques, the two that are most theoretically attractive are, the *binarized scoring rule*, proposed by [Hossain and Okui \(2013\)](#), and the *probability matching mechanism*, described by [Grether \(1992\)](#) and [Karni \(2009\)](#). However, in the context of the current paper, we viewed neither of these approaches as being preferable to the [Offerman et al. \(2009\)](#) technique, since both of these approaches introduce an additional layer of probabilities and in the study of probability bias, this is an undesirable attribute of the elicitation strategy.

¹⁵This assumes that the state-dependent payment is associated with the 100% state, not the 0% state. This assumption is made throughout the paper and the experiment.

posed in [Kothiyal et al. \(2011\)](#)) to allow for the scenario where there are state-contingent stakes (i.e. $x \neq 0$).

3 Experimental Design

The experiment was designed to test the *asymmetric updating hypothesis* using both within-subject and between-subjects comparisons of updating behavior. The experiment consisted of three treatment groups. The treatment T1.SYMMETRIC corresponds to Scenario 1 (SYMMETRIC: no exogenous state-contingent stakes) and the other two treatment groups, T2.COMBINED and T3.SEPARATE, correspond to Scenario 2 (ASYMMETRIC: state-contingent stakes) discussed above.

In T1.SYMMETRIC, one would expect the participants to be indifferent between each of the two states being realized, while in T2.COMBINED and T3.SEPARATE, the larger payment associated with state $\omega = A$ should imply that the participants prefer that this state be realized. T2.COMBINED and T3.SEPARATE are identical in terms of the financial incentives. The rationale for running two treatments with identical incentives was to conduct additional checks to ensure that our results were not driven by the influence of a hedging motive as discussed above. To do this, we varied only the way that the incentive environment was described to participants (i.e. we only varied the framing of the incentives). In T2.COMBINED, this information was summarized in a way that made it much easier for participants to notice the hedging opportunity in comparison to T3.SEPARATE.

This exogenous variation in the salience of the hedging opportunity served two purposes. Firstly, it allowed us to assess whether a hedging motive influenced the beliefs elicited. Secondly, it provided an opportunity to assess the validity of the mechanism we use to correct reported beliefs.¹⁶ The difference between these two treatments is discussed in more detail in the ‘*Incentives and Treatment Groups*’ section below.

The experiment proceeded in three stages. The first stage comprised the core belief updating task in which we elicited a sequence of reported beliefs from subjects as they received a sequence of noisy signals regarding the true state of the world, updating from an exogenously provided prior. In the second stage we collected the reported probabilities associated with known objective probabilities on the interval $[0, 1]$ required for the [Offerman et al. \(2009\)](#) correction approach, as well as data on risk preferences. In the third stage, we obtained data on several demographic characteristics as well as some further non-incentivized measures. In each of the first two stages,

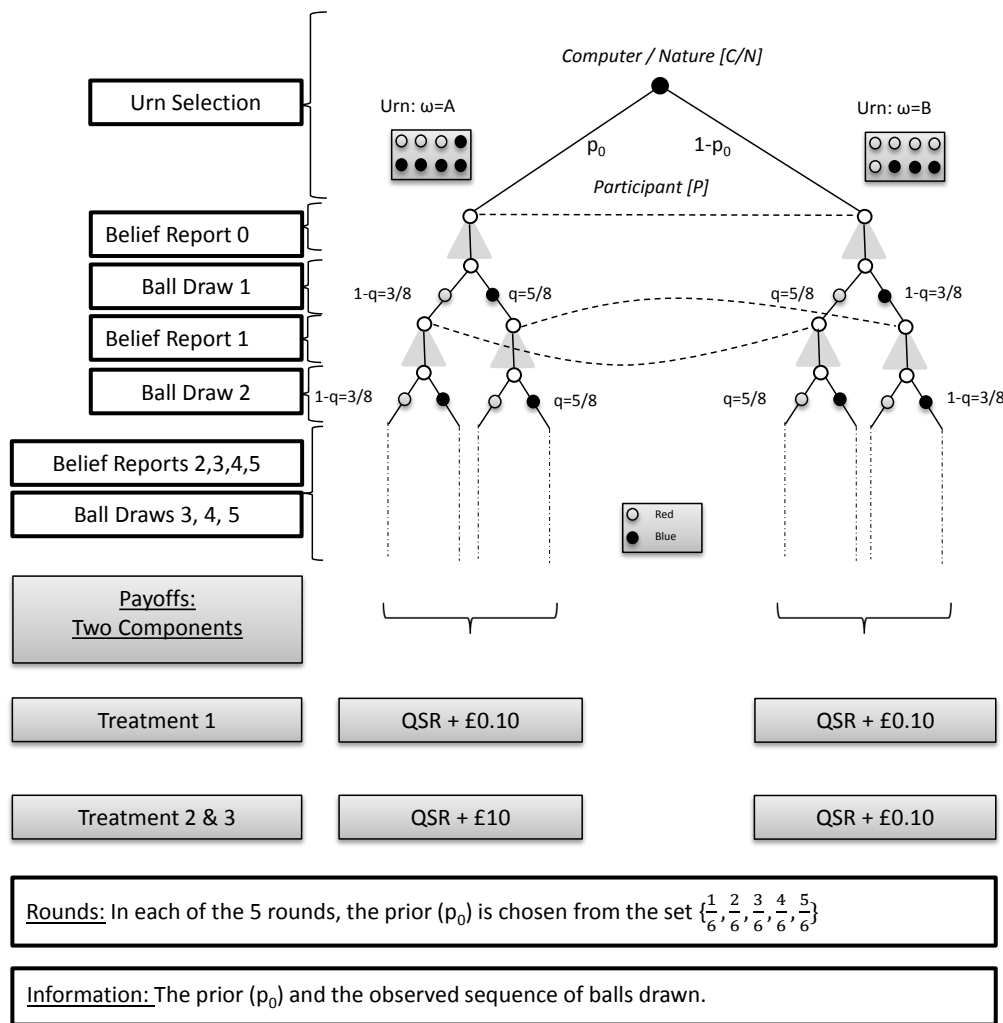
¹⁶This approach of varying the transparency of the hedging opportunity in order to evaluate the comparative static influence of the salience of hedging in practical applications in the laboratory echoes [Blanco et al. \(2010\)](#), who study whether hedging is a problem for belief elicitation in the lab. The authors conclude that it is when the hedging opportunity is salient. We will discuss the relation between our results and theirs in more detail in the Discussion section.

one of the subject's choices was chosen at random and paid out. In addition the participants received a fixed fee of £5 [€5] for completing Stage 3, as well as a show-up fee of £5 [€5].¹⁷

The Belief Updating Task (Stage 1)

The Belief Updating Task was the primary task of the experiment. We used this task to collect data about participants' belief updating behaviour. The experimental design for this task is summarized in Figure 1 and described in the following discussion.

Figure 1: Overview of Experimental Design



¹⁷In the discussion below, I will always refer only to Pounds (£), however in the experiments run at the Technical University in Berlin all the payments were made in Euros. In all cases, the payment in Pounds (£) was equivalent to the payment in Euros (€). In other words, £1 was replaced with €1. In terms of the cost of living in the two locations, this seemed more appropriate than using the actual exchange rate. In addition, it had the benefit of keeping all the quantities constant across locations.

The Belief Updating Task consisted of five rounds. In each round, participants were presented with a pair of computerized ‘urns’ containing blue and red colored balls, with each of these two urns representing one of the two states of the world. The composition of the two urns was always constant, with the state $\omega = A$ represented by the urn containing more blue balls (5 blue and 3 red), while the state $\omega = B$ was represented by the urn containing more red balls (5 red and 3 blue).

Priors

The five rounds differed from one another only in the exogenous prior probability that $\omega = A$ was the true state, with this prior, p_0 , chosen from the set $\{\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}\}$. In each round, this prior was known to the participant. The order of these rounds was randomly chosen for each individual. Conditional on the prior, p_0 , one of the two urns was then chosen through the throw of a virtual die, independently for each individual, in each round.

Belief Updating

In each round, after being informed of this prior probability, p_0 , the participant received a sequence of five partially informative signals, s_t , for $t = \{1, 2, 3, 4, 5\}$. These signals consisted of draws, with replacement, from the urn chosen for that round. Therefore, if the state of the world in a specific round was $\omega = A$ then the chance of drawing a red ball was $\frac{3}{8}$ and the chance of drawing a blue ball was $\frac{5}{8}$ for each of the draws in that round (see Figure 1).

In each round, we elicited the participant’s reported belief, r_t , about the likelihood that state $\omega = A$ was the correct state of the world, six times (i.e. for $t = \{0, 1, 2, 3, 4, 5\}$). We first elicited her reported belief, r_0 , directly after she was informed of the exogenous prior probability, p_0 , and then after she received each of her five signals we elicited r_t for $t = \{1, 2, 3, 4, 5\}$. Overall, we therefore elicited 30 reported beliefs in Stage 1 from each individual (6 reported beliefs in each of 5 rounds).¹⁸

Incentives and Treatment Groups

The Belief Updating Task was identical across treatment groups with the exception of the incentives faced by participants. In each treatment, participants’ payment consisted of two components: (i) an exogenous state-contingent payment¹⁹, and (ii) an accuracy payment that depended

¹⁸Participants were not informed about the correct urn at the end of each round. They only received feedback when their payment was calculated at the very end of the experiment.

¹⁹This was called the “urn bonus” in the experiment.

on their stated belief²⁰ and the true state (i.e. the QSR payment described in equations 6 and 7 above). In treatments T2.COMBINED and T3.SEPARATE, the state-contingent payment was substantially higher at £10 in state $\omega = A$ in comparison to £0.10 in state $\omega = B$, making $\omega = A$ the more attractive state of the world. In T1.SYMMETRIC, participants simply received an equal state-contingent payment of £0.10 in both states, $\omega = A$ and $\omega = B$, implying neither was preferable.

In all three treatments, participants received nearly identical detailed written instructions describing the belief updating task as well as the two payment components. In order to further simplify the task faced by participants and try to ensure that they understood the incentive environment they faced, we presented the QSR as a choice from a list of lotteries (this approach is also used, for example, by [Armantier and Treich \(2013\)](#) and [Offerman et al. \(2009\)](#)). To this effect, subjects were presented with *payment tables*, which informed them of the precise prospect they would face for each choice of r_t , in increments of 0.01. An abbreviated version of the three payment tables associated with each of the three treatment groups is presented in Table 2.²¹ In order to represent all payments as integers in the instructions and payment tables, we adopted the approach of using experimental points. At the end of the experiment, these experimental points were converted to money using an experimental exchange rate of 6000 points = £1.

Table 2 highlights the difference between treatments T2.COMBINED and T3.SEPARATE. While, participants in these two treatments faced precisely the same incentives, the treatments differed in terms the salience of the hedging motive. In particular, the only difference between the two treatments was the way in which the payment information was summarised in the *payment table*. As shown in Table 2, in T2.COMBINED, the *payment table* showed the combined payment from both (i) the exogenous state-contingent payment, and (ii) the accuracy payment, together. Therefore, it summarised the reduced form prospect associated with each reported probability choice (r_t) for subjects.

The motivation behind having two treatments with identical incentives, but a different presentation of the incentives, was following. The rationale for T2.COMBINED was that presenting the incentives in a combined form is the simplest and clearest way of relaying the true incentives faced to participants. The rationale for T3.SEPARATE was that if participants “narrowly bracket”, then the separate presentation of incentives could reduce the influence of the hedging motive, and therefore would induce more accurate belief reporting. By implementing both treatments, we were able to evaluate the influence of the presentation of the incentives on the reported beliefs. As we will see below, this influence was substantial and corresponds to the theoretical predictions for how a risk averse individual would act if she were hedging more when the hedging opportunity

²⁰In the experiment, we refer to the object that individuals reported about the likelihood that one of the two urns was the chosen one as their “probability judgment”. This corresponds to the reported belief in the discussion above.

²¹Note, these payment tables are abbreviated in comparison to the tables presented to participants. The only substantive difference is that Table 2 contains 21 rows, one for each 5% increment in the reported belief. In contrast, the participants received payment tables that contained 101 rows, one for each 1% increase in the reported belief.

Table 2: Comparison of Incentive Summary Tables between Treatment Groups

T1 SYMMETRIC				T2 COMBINED				T3 SEPARATE			
Probability judgment for Urn A	True Urn		Probability judgment for Urn A	True Urn		Probability judgment for Urn A	True Urn				
	Urn A	Urn B		Urn A	Urn B		Urn A	Urn B			
100	12 600	600		72 000	600	100		100	12 000	0	
95	12 570	1 770		71 970	1 770	95		95	11 970	1 170	
90	12 480	2 880		71 880	2 880	90		90	11 880	2 280	
85	12 330	3 930		71 730	3 930	85		85	11 730	3 330	
80	12 120	4 920		71 520	4 920	80		80	11 520	4 320	
75	11 850	5 850		71 250	5 850	75		75	11 250	5 250	
70	11 520	6 720		70 920	6 720	70		70	10 920	6 120	
65	11 130	7 530		70 530	7 530	65		65	10 530	6 930	
60	10 680	8 280		70 080	8 280	60		60	10 080	7 680	
55	10 170	8 970		69 570	8 970	55		55	9 570	8 370	
50	9 600	9 600		69 000	9 600	50		50	9 000	9 000	
45	8 970	10 170		68 370	10 170	45		45	8 370	9 570	
40	8 280	10 680		67 680	10 680	40		40	7 680	10 080	
35	7 530	11 130		66 120	11 130	35		35	6 930	10 530	
30	6 720	11 520		66 120	11 520	30		30	6 120	10 920	
25	5 850	11 850		65 250	11 850	25		25	5 250	11 250	
20	4 920	12 120		64 320	12 120	20		20	4 320	11 520	
15	3 930	12 330		63 330	12 330	15		15	3 330	11 730	
10	2 880	12 480		62 280	12 480	10		10	2 280	11 880	
5	1 770	12 570		61 170	12 570	5		5	1 170	11 970	
0	600	12 600		60 000	12 600	0		0	0	12 000	

was made salient.

Furthermore, an additional benefit of running both treatments was that it provided us with a way to test the internal validity of the correction mechanism we use. We will see below that, while the uncorrected distribution of beliefs differ substantially between the two treatments, the distributions of the corrected beliefs are very similar to one another.

The Offerman et al. (2009) Correction Task (Stage 2)

In the second stage of the experiment we elicited the twenty reported beliefs, r , for events with known objective probabilities required to estimate the incentive distortion function, R , for each individual. In each of the three treatments, we estimate the function R using belief from Stage 2 elicited under the same incentive environment as in the Belief Updating Task in Stage 1.

In Stage 2, participants were asked to report their probability judgment regarding the likelihood that statements of the form: “the number the computer chooses will be between 1 and 75” (i.e. $p = 0.75$), were true, after being told that the computer would randomly choose a number between 1 and 100, with each number equally probable. For T1.SYMMETRIC, this specific example of the probability of the randomly chosen number being in the interval between 1 and 75 essentially involves choosing r from the list of prospects defined by $1 - (1 - r)^2_{0.75}(1 - r^2)$. For T2.COMBINED and T3.SEPARATE, this example would involve choosing r from the list of prospects defined by $x + 1 - (1 - r)^2_{0.75}(1 - r^2)$. As in Stage 1, in each of the treatments, the Stage 2 payment table summarized the relevant payment information. For each treatment, this payment table contained identical values in Stage 1 and Stage 2. Therefore, Table 2 above also provides a summary of the Stage 2 payment tables.

The twenty reported beliefs corresponded to the objective probabilities $0.05, 0.1, \dots, 0.95$.²² At the end of the experiment, one choice from Stage 2 was randomly chosen to contribute to each participant’s final payment.

4 Data and Descriptive Evidence

The experiment was conducted at the UCL-ELSE experimental laboratory in London as well as at the WZB-TU laboratory in Berlin, with two sessions for each of the three Treatment groups at each location, making twelve sessions and 222 participants in total.²³ At both locations, participants

²²These objective probabilities were presented in a random order. Furthermore, in the likelihood statement used, such as the example in the main test, the interval given started at a randomly chosen lower bound (from the feasible set of lower bounds for that specific objective probability). Nineteen of the twenty reported beliefs were unique, and the twentieth was a randomly chosen repetition of one of the first nineteen.

²³The experiment was therefore completely symmetric between the two locations, with approximately the same number of participants in each treatment group from each location.

were solicited through an online database using ORSEE (Greiner, 2015) and the experiment was run using the experimental software, z-Tree (Fischbacher, 2007). On average, sessions lasted approximately 1.5 hours and the average participant earned £19.7 in London and €20.3 in Berlin. Realized payments ranged between £11 [€11] and £34 [€34].

One challenge faced by belief updating studies is ensuring that subjects understand and engage with the task. In order to facilitate this, we were careful to ensure that the instructions received were as clear and simple as possible. Nonetheless, there remained a non-trivial fraction of participants who took decisions to update in the ‘incorrect’ direction²⁴ upon receiving new information. In order to ensure that the behavior we are studying is reflective of actual updating behavior of individuals who understood and engaged with the task, we restrict our sample for our main analysis by removing rounds where an individual updates in the incorrect direction.²⁵ However, we also estimate all the main results on the full sample, and the general patterns of behavior are similar.

While randomization to treatment group should ensure that the samples are balanced on observable and unobservable characteristics, Table 9 provides a check that the selection of our preferred sample has not substantially biased our treatments groups by reporting the sample means of a set of individual characteristics for each treatment group. Overall, the treatments appear to be balanced, with the exception that individuals in T3.SEPARATE are more likely to speak English at home than individuals in T2.COMBINED.

5 Empirical Specification

In this section, we first discuss the calibration exercise used to correct the reported beliefs. Second, we describe the core estimation equations used in our analysis. These build on the work by Möbius et al. (2014) as their data has a similar structure to ours. One key difference in our data is the exogenous assignment of the participants’ entire information set. We exploit this feature of our data to address endogeneity issues that can arise when studying belief updating.

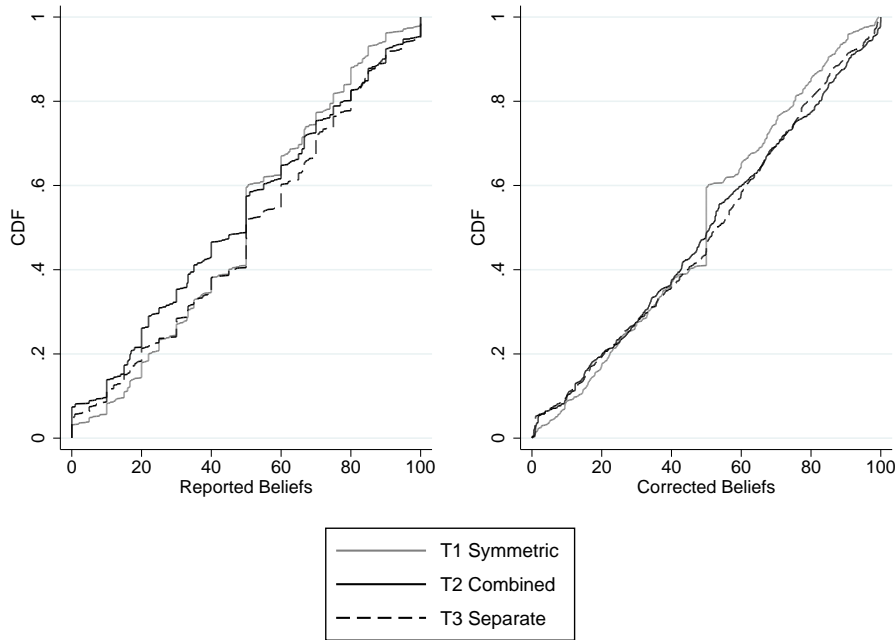
²⁴‘Incorrect’ in the sense that the participant received a signal that increased the likelihood that one state was the true state, but the individual updated her reported belief in the opposite direction. For example, if the individual drew a red ball, but increased her reported belief regarding the probability that the state was $\omega = A$, then we would label this as an update in the incorrect direction.

²⁵More specifically, for our preferred sample, we remove all individuals who make more than five out of twenty five updating decisions in the incorrect direction. This comprised eighteen percent of our subjects. This is comparable to the twenty five percent of individuals who make at least one mistake out of four decisions in Möbius et al. (2014). For those who updated in the wrong direction five or fewer times, we remove only the round of decisions in which they made a mistake. This removes 210 out of the remaining 910 rounds. We estimate our results on the full sample and all the main results remain the same.

The Belief Correction Procedure

The belief correction procedure that we adopt involves assuming a flexible parametric form for the participants' utility and probability weighting functions in order to estimate their personal belief-distortion function (this is the R function discussed in Equation 8 in Appendix B.2.1). We estimate this function for each individual separately in order to correct the reported beliefs at the individual level. Essentially, we are simply fitting a curve through each subject's belief elicitation distortion, for the relevant incentive environment. Appendix B.2.4 contains a more detailed discussion of the estimation of the belief correction function, and illustrates the main ideas through Figures 5 and 6. Figure 5 displays the average correction curves for each of the treatment groups across all individuals. However, at the individual level, there is a large degree of heterogeneity in the degree to which individuals distorted their reported belief away from their actual belief, given the incentive environment. Figure 6 displays the correction curves estimated for two individuals from each treatment group.

Figure 2: CDFs of Corrected and Uncorrected Reported Beliefs



In Figure 2 we plot the CDFs of the uncorrected reported beliefs as well as the reported beliefs that have been corrected at the individual level. The left panel displays the distribution of reported beliefs, prior to correction, in each of the three treatments. One interesting feature of this figure is that, in spite of the fact that subjects in T2.COMBINED and T3.SEPARATE are offered precisely the same incentives, the distributions of reported beliefs of these two groups are significantly different from one another (Mann-Whitney rank-sum test, $p < 0.01$). The larger mass of reported

beliefs to the left of 50 in T2.COMBINED suggests that individuals are more likely to respond to the hedging opportunity when it is more salient.

This right panel shows the corrected beliefs. It suggests that the belief correction approach that we use was successful in removing the strong hedging influence in the T2.COMBINED treatment since the correction procedure removes the majority of the difference between the distributions of beliefs between the two treatments with identical incentives, T2.COMBINED and T3.SEPARATE. This evidence underscores the usefulness of the belief correction mechanism in reducing the bias in the reported beliefs, when there is a salient hedging opportunity.

With the corrected beliefs in hand, we can proceed to the main analysis of belief updating in our sample. The analysis is done using both the corrected and uncorrected beliefs.²⁶

Core Estimation Specifications

Our core estimation equations aim to test for systematic patterns in updating behavior, within the framework developed in Equation 1. Firstly, we examine whether there are systematic deviations from Bayes' rule in updating behavior, independent of having a stake in one of the two states of the world. Secondly, we assess the influence that having a stake in one of the two states of the world has by (i) testing for an asymmetry in updating *within* treatments where there is a state-contingent stake (i.e. T2.COMBINED and T3.SEPARATE); and (ii) testing whether there are differences in updating behavior *between* the treatments with and without a state-contingent stake.

The first estimation equation follows directly from Equation 1, allowing us to test the asymmetric updating hypothesis, and also to test for other common deviations from Bayes' rule, such as a *confirmatory bias* or *base rate neglect*²⁷:

$$\tilde{\pi}_{i,j,t+1} = \delta \tilde{\pi}_{i,j,t} + \gamma_a \tilde{q} \cdot 1(s_{i,j,t+1} = a) - \gamma_b \tilde{q} \cdot 1(s_{i,j,t+1} = b) + \epsilon_{i,j,t+1} \quad (2)$$

where $\tilde{\pi}_{i,j,t} = \text{logit}(\pi_{i,j,t})$ and $\tilde{q} = \log(\frac{q}{1-q})$; j refers to a round of decisions; and the errors, ϵ_{ijt} , are clustered at the individual (i) level.

²⁶The analysis using the uncorrected beliefs is essentially an analysis using beliefs elicited using the QSR, which is by far the most common approach used in the belief elicitation literature to date. Therefore, these results have greater comparability with much of the experimental literature. However, in line with the discussion above, our preferred specification uses the corrected beliefs as these should be closer to the subjects' true beliefs.

²⁷Here, base rate neglect is the tendency to underweight or partially disregard one's prior regarding which urn is more likely to be generating the signal. This implies that when the individual receives a new signal, she weights information that supports her prior less heavily than information that opposes it, relative to a Bayesian. In this case, when one's prior regarding state $\omega = A$ is greater than 0.5, $\pi_{t-1} > 0.5$, a participant who is prone to *base rate neglect* weights signals that support state $\omega = A$ less heavily than signals that support state $\omega = B$; and vice versa when her prior is $\pi_{t-1} < 0.5$, using Bayesian updating as the benchmark.

In order to determine the belief updating pattern *within* each incentive environment, we estimate this equation separately for each treatment; and then to test for significant differences *between* the coefficients in different incentive environments, we pool our sample and interact the treatment variable with all three of the coefficients in this equation (i.e. δ, γ_a , and γ_b). This provides us with a test of whether the parameters differ between either of the two ASYMMETRIC treatments and the SYMMETRIC treatment.

Endogeneity of the Lagged Belief

One potential concern with the identification of the parameters of Equation 2 is the common issue where the right hand side contains lagged versions of the dependent variable. This implies that there is a possible endogeneity of the lagged beliefs, $\pi_{i,j,t}$, if they are correlated with the error term, $E\{\tilde{\pi}_{i,j,t} \epsilon_{i,j,t+1}\} \neq 0$.²⁸ If this is the case, it can result in biased and inconsistent estimators for the parameters of the regression. Our experiment was designed to avoid this issue by virtue of exogenously assigning the subjects' entire information set. This allows us to use the exogenously assigned prior probability of state $\omega = A$ being the true state, $p_{i,j,t=0}$, as well as the sequence of signals observed, s_t , to construct an instrument for the lagged belief, $\pi_{i,j,t}$, in Specification 2.²⁹ We do this by calculating the objective Bayesian posterior, given the agent's information set at time t , and using this as an instrument for her belief, $\pi_{i,j,t}$.

The approach used here also avoids a second type of endogeneity issue that can arise when studying belief updating when the states of the world are a functions of *personal characteristics* (e.g. when examining beliefs regarding individual attributes, such as one's own skills, IQ, or beauty) or *personal choices*. When this is the case, the conditional probability of observing a specific signal depends on the state of the world, and therefore can be correlated with personal characteristics.³⁰

²⁸For example, this would be the case if there is individual heterogeneity in the way individuals respond to information. We provide evidence below that this individual heterogeneity is present.

²⁹More precisely, we instrument using the logit of the accurate Bayesian posterior, $\text{logit}(p_{i,j,t})$, as an instrument for the logit of the lagged belief, $\tilde{\pi}_{i,j,t} = \text{logit}(\pi_{i,j,t})$ in Specification 2.

³⁰Consider the following toy example to illustrate this. Assume there are two states of the world—being TALL and SHORT—and being tall is viewed as “good”. Individuals are initially uncertain about their height group (because only relative height is important). Now, assume that TALL individuals are more responsive to new information than SHORT individuals. For example, assume TALL individuals update exactly according to Bayes' rule, but SHORT individuals don't update their prior beliefs at all in response to new information. Now, let subjects receive a sequence of noisy signals that are informative (in the sense that, for each type, the signal they receive is correct more than fifty percent of the time). In this toy example, both types respond symmetrically to ‘good-news’ and ‘bad-news’, but because TALL individuals are always more likely to receive ‘good-news’, and SHORT individuals are always more likely to receive ‘bad-news’, the resulting aggregate behavior will appear to display the “good-news, bad news effect”. This is due to the fact that, on average, a good signal was more likely received by a TALL individual, who updates like a Bayesian, and a bad signal is more likely to go to a SHORT individual who doesn't update at all. Furthermore, updating would display conservatism on average. See Appendix C for further details.

6 Results

Table 3 reports the results from estimating Equation 2 for each of the treatment groups separately. These estimates provide an overview of the updating behavior of the average individual within each of the three treatment groups. Within each treatment group, we report the results for both the OLS (top panel) and the IV (bottom panel) estimates discussed above. Columns (1a), (2a) and (3a) use the uncorrected reported beliefs, while columns (1b), (2b) and (3b) use the corrected beliefs. Every coefficient in the table is statistically different from 0 at the 1% level. Since our primary interest is in testing whether the coefficients are different from 1, in this table, we use asterisks to reflect the significance of a t-test of whether a coefficient is statistically different from 1.

Perhaps the most striking features of this table are: (i) the similarity in the updating patterns across the three treatment groups; and (ii) that for the average individual, the observed updating behavior is close to Bayesian in all three treatment groups. The p-values from the test of the null hypothesis, $H_0 : \gamma_a = \gamma_b$, show that in none of the three treatment groups do we observe a statistically significant difference (at the 5% level) between the responsiveness to the signals in favor of $\omega = A$ and $\omega = B$ (i.e. we don't observe asymmetric updating).

Both the OLS results in the top panel, and the IV results in the bottom panel, indicate that the responsiveness to new information was, on average, not statistically different to that of a Bayesian, since both γ_a and γ_b are not significantly different to 1 at the 5 percent level. The primary difference between the OLS results and the IV results is that, while the OLS estimates suggest a small degree of base rate neglect across all three treatments ($\delta < 1$), once we control for the possible sources of endogeneity discussed above using our instrumental variable strategy, the estimates are no longer indicative of base rate neglect. Since the OLS estimates may be biased³¹, the IV estimates represent our preferred results. The first stage regression results for the IV estimation are reported in the Appendices in Table 8, indicating that we don't have a weak instrument issue. Importantly, however, the form of endogeneity addressed by IV estimation only pertains to potential biases in the δ parameter, since it addresses endogeneity in the prior belief variable. It is therefore reassuring that for the estimates of our primary parameters of interest, namely γ_a and γ_b , the estimates are largely consistent across all the estimation specifications reported in Table 3. A reason for this is that the signals, $s_{i,j,t+1}$, that subjects receive are always completely exogenous, both across rounds, and with respect to the subject's personal characteristics. Furthermore, the distribution of signals observed is also exogenous, and balanced in expectation. This helps to

³¹As mentioned above, one candidate for biased estimates is due to the fact that the posterior belief at time t is a left-hand-side variable in time t , but also serves as the right-hand-side variable through being the prior belief at time $t + 1$. This can lead to biased estimates if there is individual heterogeneity in updating. Our instrumental variable strategy uses the exogeneity of the priors and signals that subjects have been endowed with (summarized as the objective Bayesian probability at that point in time) to instrument for their prior beliefs, thereby removing this endogeneity of the prior beliefs in the regressions.

avoid other sources of endogeneity (see, e.g., Appendix C) and to alleviate the influence of other potential confounding belief updating biases (see the Discussion section below).

Table 3: Average Updating Behavior across Treatments

	<u>T1: SYMMETRIC</u>		<u>T2: COMBINED</u>		<u>T3: SEPARATE</u>	
	Reported (1a)	Corrected (1b)	Reported (2a)	Corrected (2b)	Reported (3a)	Corrected (3b)
<u>OLS</u>						
δ	0.90 (0.03)***	0.90 (0.03)***	0.86 (0.04)***	0.86 (0.04)***	0.91 (0.03)***	0.93 (0.02)***
γ_a	1.12 (0.12)	1.09 (0.11)	1.05 (0.12)	1.06 (0.12)	1.25 (0.14)*	1.16 (0.11)
γ_b	1.21 (0.13)*	1.17 (0.11)	1.19 (0.15)	1.12 (0.13)	1.20 (0.12)	1.13 (0.10)
$p(H_0 : \gamma_a = \gamma_b)$	0.30	0.32	0.20	0.48	0.58	0.73
N	1,075	1,075	1,285	1,285	1,140	1,140
R^2	0.71	0.73	0.73	0.74	0.81	0.84
<u>IV</u>						
δ	0.99 (0.03)	0.99 (0.03)	1.00 (0.02)	0.99 (0.02)	0.99 (0.03)	0.99 (0.02)
γ_a	1.12 (0.12)	1.09 (0.11)	1.08 (0.12)	1.02 (0.11)	1.22 (0.13)*	1.14 (0.11)
γ_b	1.21 (0.12)*	1.16 (0.11)	1.11 (0.13)	1.12 (0.12)	1.18 (0.12)	1.11 (0.10)
$p(H_0 : \gamma_a = \gamma_b)$	0.30	0.31	0.76	0.25	0.63	0.74
N	1,075	1,075	1,285	1,285	1,140	1,140
1st Stage F	61.38	84.04	107.98	107.01	76.90	95.45

(i) Standard errors in parentheses (clustered at the individual level).

(ii) All coefficients are significantly different from 0 at the 1% level. Therefore, t-tests of the null hypothesis (H_0 : Coefficient = 1) are reported: * = 10%, ** = 5%, *** = 1%.

(iii) The rows corresponding to $p(H_0 : \gamma_a = \gamma_b)$ report the p-statistic from a t-test of the equality of the coefficients γ_a and γ_b (i.e. a test of the asymmetric updating hypothesis).

It is worth noticing that although we observe a substantial difference in the levels of the corrected and uncorrected beliefs in Figure 2 above, the estimates for updating in Table 3 are quite similar for the corrected and uncorrected beliefs. One explanation for this apparent inconsistency is the following. If the degree of hedging by an individual is similar for both the prior and posterior belief, then the correction could have a sizeable effect on the levels of both beliefs, but not result in

the large difference in the estimated updating parameters, since updating pertains to the change in the belief rather than the level.

A Model Free Test of the Asymmetric Updating Hypothesis

In order to alleviate the potential concern that these results are dependent on the functional form of our empirical specification, we conduct a model-free test of the *asymmetric updating hypothesis*. Perhaps the simplest and most direct test of this hypothesis is obtained by directly comparing the posterior beliefs formed in two scenarios where the information set is identical, but the rewards associated with one of the states of the world is varied. Our data is well suited for conducting this exercise.

We do this by considering a comparison of information-set-equivalent posterior beliefs after individuals have received only (i) the exogenous prior and (ii) a single ball draw.³² This allows us to test the *asymmetric updating hypothesis* while remaining agnostic regarding the process that guides belief updating, testing only whether it is symmetric. Our data allows us to conduct two comparisons of information-set-equivalent posterior beliefs—a *within-subject* and a *between-subject* comparison.

Firstly, we can compare posterior beliefs, π_1 , formed with identical information sets $\{p_0, s_1\}$ *between* treatment groups, where the payments associated with states of the world differ. For example, we can compare the average posterior formed after an identical prior, e.g. $p_0 = \frac{1}{6}$, and an identical signal, e.g. $s_1 = a$ (i.e. a blue ball), across treatments.

Secondly, we can compare information-set-equivalent posterior beliefs *within* treatment groups. This comparison involves comparing π_1 after $\{p_0 = p, s_1 = s\}$ with $1 - \pi_1$ after $\{p_0 = 1 - p, s_1 = s^c\}$ where s^c is the complementary signal to s .³³ For example, we can compare the posterior, π_1 , formed after a prior of $p_0 = \frac{1}{6}$ and the signal $s_1 = a$ (i.e. a blue ball), with $1 - \pi_1$ after a prior of $p_0 = \frac{5}{6}$ and the signal $s_1 = b$ (i.e. a red ball). To see why this comparison involves a comparison of information-set-equivalent posterior beliefs, recall that the experiment is designed to be completely symmetric in terms of information, with the information content of a red ball exactly the same as a blue ball, except in support of the other state of the world. Therefore, if an individual updates symmetrically, then $\pi_1|_{\{p_0=p, s_1=s\}} = 1 - \pi_1|_{\{p_0=1-p, s_1=s^c\}}$. This prediction does not rely on Bayes' rule (although it is an implication of Bayes' rule), but rather only requires symmetric updating, and therefore it provides us with a non-parametric test of the *asymmetric updating hypothesis*.

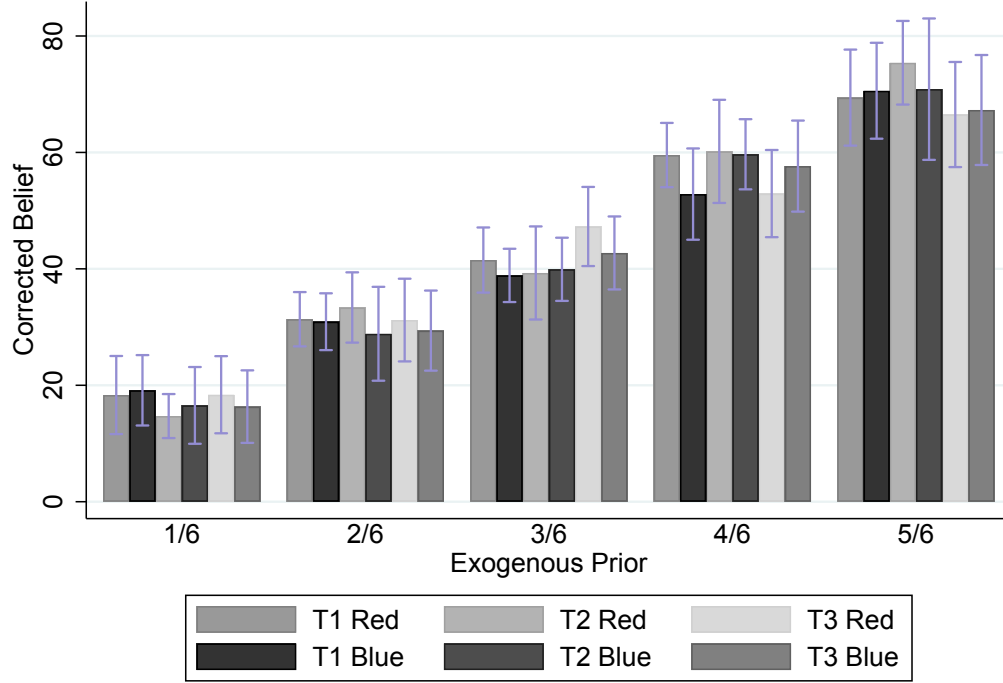
Figure 3 depicts both of these comparisons, with each group of six bars collecting together the relevant information-set-equivalent groups. Each bar presents the mean posterior belief for that

³²i.e. the information set consists of the exogenously assigned prior, p_0 , and a single signal, $s_1 \in \{a, b\}$.

³³Therefore, if $s = a$ then $s^c = b$ and vice versa.

group, as well as a 95% confidence interval around the mean.³⁴ Each group is labeled on the x-axis by the prior belief associated with the ‘red’ bars, which correspond to the information sets that include a red ball as a signal (i.e. $s_1 = b$). The ‘blue’ bars report the mean of $1 - \pi_1$, for information sets containing a blue ball (i.e. $s_1 = a$) and for these bars the x-axis label corresponds to $1 - p_0$. Within each group, the first two bars represent the average posterior beliefs in T1.SYMMETRIC; the second pair of bars depict the same for T2.COMBINED; and the third pair of bars for T3.SEPARATE.

Figure 3: Comparison of Beliefs after the Receipt of a Single Signal and an Exogenous Prior



Notes: (i) Bar denotes 95 percent confidence interval around the mean; (ii) Beliefs after blue balls reflect $(100 - \text{belief})$.

The results displayed in Figure 3 show that there are no systematic differences between posterior beliefs within information-set-equivalent groups, neither *within* nor *between* treatment groups. Furthermore, when testing non-parametrically whether there are differences *within* or *between* treatment groups for information-set-equivalent groups, none of the 45 relevant binary comparisons³⁵ are significant at the 5 percent significance level under a Mann-Whitney test, suggesting that we cannot reject the hypothesis that the posterior beliefs within information-set-equivalent groups are drawn from the same distribution. This lends support to the results described above which indicate that we fail to find evidence in support of the *asymmetric updating hypothesis*.

³⁴The calculation of these confidence intervals assumes that the data are distributed according to a t-distribution.

³⁵For these comparisons, for each exogenous prior, we test the following binary comparisons: (i) within treatment group, we test between those that received the $s_1 = a$ and $s_1 = b$ signals (3×5 binary comparisons); (ii) for those that received the same signal, s_1 , we test between treatment groups (6×5 binary comparisons).

Robustness Exercises

In addition to this model-free test, to check for the robustness of the belief updating results for the average individual presented in Table 3, we conducted several robustness exercises. These exercises, and their corresponding results, are discussed in detail in Appendix A.

The first exercise examines whether the results from the main specification described in Equation 2 are robust to first differencing the dependent variable (i.e. this considers how new information shifts the change in beliefs, imposing the assumption that $\delta = 1$). The second subsection extends the main empirical specification to allow for individual-specific updating parameters. For both of these specifications, an ex post power analysis is conducted, reporting the MDE for a significance level of $\alpha = 0.05$ and a power of $\kappa = 0.8$. The third subsection pools all the observations across the three treatments together, and then tests whether the average updating parameters differ across treatments by interacting treatment group dummies with the regressors of the main specification described in Equation 2.

The results from all of these exercises are highly consistent with those in Table 3 and fail to provide any evidence in favor of an asymmetry in updating.

Heterogeneity in Updating Behavior

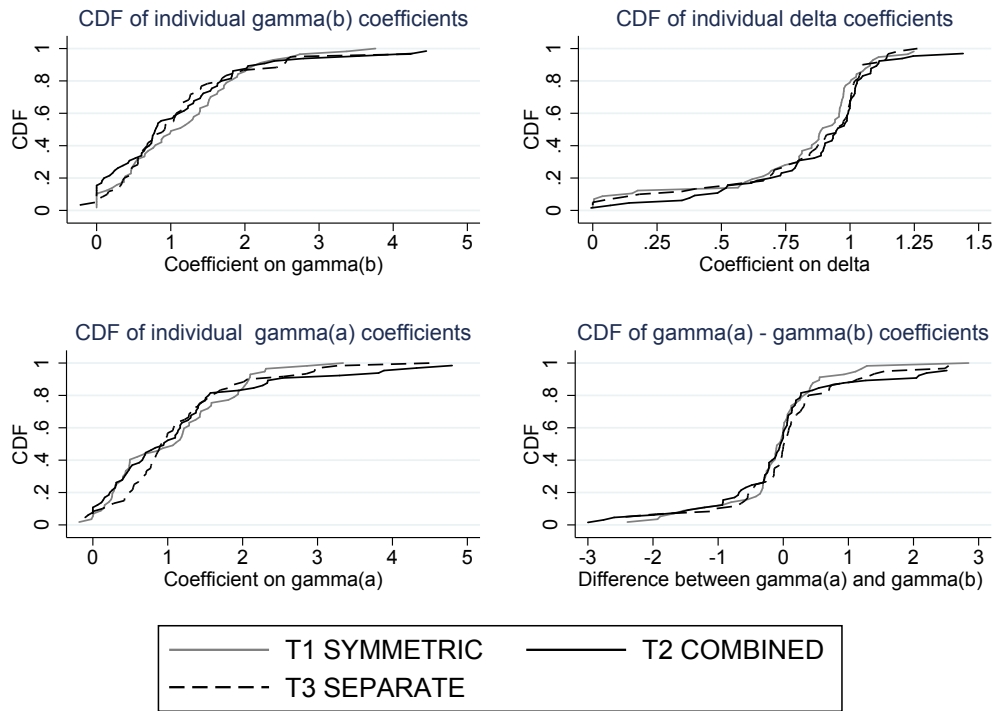
In order to investigate whether the aggregate results are masking heterogeneity in updating behavior, we estimate Specification 2 at the individual level and collect the parameters. The distributions of these individual level parameters are reported in Figure 4. Perhaps the most conspicuous feature of this figure is the fact that all three treatment groups display such similar parameter distributions in each of the panels - i.e. for each of the parameters. Testing for differences between the underlying distributions from which the parameters are drawn in the different treatment groups fails to detect any statistically significant differences in any of the four panels.³⁶ The upper-right panel shows that the majority of individuals have an estimated δ parameter in the interval $[0.6, 1.1]$, with a large proportion of these concentrated around 1 in all three treatment groups. The two left-hand panels show that there is substantially more individual heterogeneity in the estimated γ_a and γ_b parameters, which are dispersed over the interval $[0, 3.5]$ in all three treatments.

With such a large degree of variation in the individual level parameters, a natural conjecture to make is that, while we do not observe asymmetric updating at the aggregate level, it is entirely

³⁶More specifically, for each of the four panels in Figure 4, we conducted three binary Mann-Whitney tests comparing each possible pair of treatments. In total, these twelve statistical tests failed to detect any significant difference in the underlying parameter distributions between the treatment groups at the 10% level. In addition, another twenty-four similar tests for differences in the mean (t-test) and median (chi-squared) of the underlying distributions between treatments fail to detect any statistically significant differences at the 10% level.

plausible that there may be a subsample of individuals who are optimistic updaters and another subsample of individuals who are pessimistic updaters. If these two subsamples are of a similar size and their bias is of a similar magnitude, we would observe no asymmetry at the aggregate level. The lower-right panel of Figure 4 suggests that this is not the case by plotting the distribution of the individual level difference between the γ_a and γ_b parameters. The majority of the distribution is concentrated in a narrow interval around 0 for all three treatment groups, suggesting that there is no asymmetry for any sizable subsample. Furthermore, this conclusion is supported by the fact there are no significant differences between the distributions of updating parameters observed across the three treatments in any of the four panels, since the motive for a ‘good-news, bad-news’ effect is switched off in the T1.SYMMETRIC treatment.

Figure 4: Distributions of Individual Level Updating Parameters.



7 Discussion

7.1 Heterogeneity in results observed in the asymmetric updating literature

A central question that emerges from the discussion above is why we observe no evidence of a ‘good-news, bad-news’ effect here, while some other influential contributions to this literature

have found evidence for such an effect. In his excellent chapter in the Handbook of Behavioral Economics, Benjamin (2019) points out that, more generally, the evidence in this nascent literature is so far very mixed. In the economics literature, three papers find evidence in favor of stronger inference from *good news*: Eil and Rao (2011), Möbius et al. (2014) and Charness and Dave (2017).³⁷ In contrast, there are three papers that find evidence of stronger inference from *bad news*: Ertac (2011), Kuhnen (2015) and Coutts (2019).³⁸ Furthermore, in addition to the current paper, there are four other papers that find no evidence in favor of a preference-biased asymmetry in belief updating: Grossman and Owens (2012), Schwardmann and Van der Weele (2016), Gotthard-Real (2017) and Buser et al. (2018).³⁹ In the psychology literature, however, there appears to be a near-consensus arguing that there does exist an asymmetry in belief updating in favor of good news (see, e.g., Sharot et al. (2011), Sharot et al. (2012), Kuzmanovic et al. (2015), and Marks and Baines (2017), amongst others). A notable exception is provided by Shah et al. (2016) who argue that many of the contributions to this literature suffer from methodological concerns; Garrett and Sharot (2017) offer a rebuttal, claiming that optimistically biased updating is robust to these concerns.

So far, there seems to be no clear pattern organizing these heterogeneous results. However, below I offer a discussion of some of the candidate explanations for the heterogeneity in observed results. In general, the candidate explanations for these heterogeneous results fall into two categories: (1) the hypothesis that contextual factors mediate asymmetric updating: belief updating is influenced by preferences, but this preference-biased updating is switched on or off by contextual factors. (2) the hypothesis that asymmetric updating is sometimes misidentified: belief updating is *not* actually influenced by preferences, but rather, what appears to be asymmetric updating is driven by a different cognitive bias (e.g. prior-biased inference). The majority of the explanations discussed below fall into the first category.

7.1.1 Information structure

One avenue of enquiry for attempting to reconcile the results is to consider the differences in the information structures across experiments. For example, while several of the studies adopt a two-

³⁷Mayraz (2013) also presents evidence in favor of individuals forming motivated beliefs that are distorted towards more desirable states, however in his experiment one cannot calculate the Bayesian posterior, so it is less comparable to the other studies in this literature.

³⁸Kuhnen (2015) differs slightly from the other papers in this literature by studying investor learning from financial information in the domain of gains versus losses (the experimental game is framed as a financial decision). She finds that investors are especially reactive to bad outcomes in the loss domain, implying that she observes negative asymmetric updating in the loss domain (and not the gain domain), and overall more pessimistic beliefs in the loss domain. The implication is that individuals learn differently in the domain of losses in comparison to the domain of gains.

³⁹Additionally, while Eil and Rao (2011) found evidence of an asymmetry in favor of good-news in the domain of beliefs about one's own *Beauty*, they did not find evidence of an asymmetry in the domain of beliefs about one's own *IQ*.

state bookbag-and-poker-chip experimental paradigm, [Ertac \(2011\)](#) uses a three-state structure with signals that are perfectly informative about one state, and [Eil and Rao \(2011\)](#) consider a ten-state updating task with binary signals. However, this does not seem to be driving the differences in results, since we observe heterogeneous results amongst papers with similar information structures—e.g. restricting attention only to the papers with two-state structures with binary signals (e.g. the current paper, [Möbius et al. \(2014\)](#), [Gotthard-Real \(2017\)](#), [Coutts \(2019\)](#)) yields mixed results.

7.1.2 Priors

Focusing only on two-state experiments, there is substantial variation in average prior belief across experiments, with [Coutts \(2019\)](#) (by design) observing relatively low average priors in comparison to [Möbius et al. \(2014\)](#), for example. If belief updating is influenced by prior beliefs (e.g. a confirmatory bias), then what looks like preference-biased belief updating may be driven by a completely different cognitive deviation from Bayesian updating, namely prior-biased updating. However, if we look at the papers that find evidence for preference-biased updating, [Charness and Dave \(2017\)](#) do find evidence in favor of prior-biased updating, while [Eil and Rao \(2011\)](#) and [Möbius et al. \(2014\)](#) do not find evidence of prior-biased updating.⁴⁰ This speaks against the explanation that what appears to be asymmetric belief updating due to preferences is actually driven by a confirmatory bias.⁴¹

7.1.3 Ambiguity

One important dimension that differentiates the current paper from much of the literature is that in our experiment subjects are exogenously endowed with accurate point estimate priors, as opposed to updating from subjectively formed prior beliefs.⁴² There are advantages and disadvantages to this approach. It brings increased experimental control and improved causal identification, but this comes at the expense of reduced realism and perhaps a slightly less natural setting.⁴³ However, this discussion highlights a key assumption that is typically made in this liter-

⁴⁰I refer the interested reader to [Benjamin \(2019\)](#) Section 8 for a more detailed discussion of prior-biased updating.

⁴¹A second possible way in which priors can be related to belief updating is through the distribution of signals observed. For example, compare an individual with a low prior with an individual with a high prior in a typical two-state experimental framework. The individual with the low prior is also more likely to be in the low state (assuming beliefs are correlated with the true states). Therefore, in these two-state bookbag-and-poker-chip experiments, the low prior belief individual is more likely to receive more low signals. This implies that if there is any deviation from Bayesian updating that is related to the distribution of signals observed, then this can look like a preference-biased asymmetry in updating when the average prior in the experiment is different from 0.5. This point is discussed further in [Coutts \(2019\)](#). In the current paper, priors and signals are balanced for each individual, alleviating this concern.

⁴²An exception to this is [Gotthard-Real \(2017\)](#), who uses a similar experimental design to the current paper, also endowing participants with an exogenous prior point belief.

⁴³An interesting recent contribution to the belief updating literature by [Le Yaouanq and Schwardmann \(2019\)](#) proposes a methodology for studying belief updating in more natural settings, while still maintaining experimental

ature, namely that subjects are probabilistically sophisticated and therefore have in mind a point estimate of this probability (or update as if they hold a point estimate prior). In cases where subjects must form their own subjective prior belief, this assumption is not innocuous. If, instead, the beliefs subjects hold are ambiguous⁴⁴, then simple Bayesian updating of a single point estimate may no longer be the most appropriate (normative or descriptive) benchmark. Firstly, there are several competing theoretical models of belief updating in the presence of ambiguity with differing predictions, e.g. *full Bayesian updating* (Jaffray, 1989; Pires, 2002) and *maximum likelihood updating* (Gilboa and Schmeidler, 1993), and some recent experimental evidence testing between them (Ngangoue, 2018; Liang, 2019). Secondly, one might postulate that there is greater scope for motivated reasoning when one is updating beliefs from ambiguous priors (or ambiguous signals) in comparison to belief updating from exogenously endowed point estimate priors (and signals with clearly defined informativeness).

However, the existing evidence suggests that the presence or absence of ambiguity is not the primary explanation for the differing results. Even within the set of papers with home-grown subjective priors (e.g. Eil and Rao (2011), Ertac (2011), Möbius et al. (2014), and Coutts (2019)), the results are very mixed.

7.1.4 Domain of belief updating

Typically, we treat belief formation as being domain-independent. However, it seems natural to consider the possibility that humans evolved to process information about their *physical environment* differently from information about their *self* and their *social environment*. For example, the mental processes involved in forming a belief about the likelihood of future rainfall may be fundamentally different to those involved in forming a strategic belief about the probability that another individual will be trustworthy in a specific scenario. The latter may involve forming a mental model of the other individual's incentives and personal characteristics, and mapping them into the specific scenario. Some papers in this literature have explored this question by asking whether we update differently about a given fundamental characteristic of one's own self in comparison to the same fundamental characteristic of another individual (e.g., Möbius et al. (2014) and Coutts (2019)). Furthermore, recent theoretical and experimental work has studied how individuals attribute outcomes to their *self* versus an external fundamental from their *physical* or *social environment* (see, e.g., Heidhues et al. (2018), Hestermann and Le Yaouanq (2018), and

control and permitting a comparison with Bayesian updating. Their approach is attractive, since it facilitates the study of belief updating of home-grown priors upon receipt unstructured information that is unobserved by the experimenter.

⁴⁴For example, in a typical two-state experiment, I might believe that my chance of being in the top half of an ability distribution is somewhere between 50 and 70 percent. Even if I don't have an ambiguous belief, but rather hold the belief that my chance of being in the top half is uniformly distributed between 50 and 70, there is evidence that attitudes towards compound lotteries is closely related to attitudes towards ambiguity. (Halevy, 2007)

Coutts et al. (2019)).

The influence of the domain on belief updating may matter for the asymmetric updating literature because this literature as a whole considers updating scenarios pertaining to both the *environment* (with ‘good’ states typically represented by high monetary payments) and to the *self* (where ‘good’ states pertain to a desirable individual characteristic).⁴⁵ One might posit that asymmetric updating only manifests in certain domains.⁴⁶ However, even within the group of studies considering beliefs about the self, the evidence is mixed (e.g. Eil and Rao (2011), Ertac (2011), Möbius et al. (2014), and Coutts (2019)).

7.1.5 Outcomes and stake size

One caveat to the results reported in the current paper is the financial stakes in play are not extremely large, and therefore it is feasible that the results would change in the presence of larger stakes. While small stake sizes are a standard caveat to most laboratory experiments, it is perhaps a larger concern here, since the asymmetric updating hypothesis may rely on the degree to which individuals desire the occurrence of the good state of the world. However, Coutts (2019) investigates the role played by stakes, increasing the stake size up to \$80 and finds no evidence that it plays a role. This suggests that stake size may not be a pivotal concern.⁴⁷ Nonetheless, it is worth keeping this caveat in mind when interpreting the results.

Overall, the results in this literature appear to be somewhat incongruous, and there doesn’t yet seem to be a compelling explanation for the observed pattern of behavior observed across studies. As suggested by Daniel Benjamin (2019), resolving these seeming contradictory results should be a priority. The current paper adds an incremental contribution towards this pursuit by demonstrating the absence of a good-news, bad-news effect in one cell of this contextual matrix—namely, a context with financial outcomes, minimal ambiguity, and a minimal role for the individual’s ego.

⁴⁵A closely related, but slightly different, taxonomy of the domain space considered in this literature is obtained by distinguishing domains where an individual’s ego is present from domains in which it is absent. However, for most of the discussion here, this overlaps closely with the self vs physical environment dichotomy.

⁴⁶One potential explanation for the difference in belief updating between the domains of self-image and financial decision making is the idea that ego maintenance could yield evolutionary benefits. In particular, a positive asymmetry in updating about one’s self-image would lead to overconfident beliefs, and several authors have posited that maintaining a high self-confidence may be associated with evolutionary advantages (see, e.g., Bernardo and Welch (2001); Heifetz et al. (2007); Johnson and Fowler (2011); Burks et al. (2013); Schwardmann and Van der Weele (2016)). In contrast, asymmetric updating about external states of the world would lead to overoptimism which is likely to lead to costly mistakes.

⁴⁷One might argue that even this large monetary stake size is small in comparison to how much individuals care about their *self*. However, the heterogeneity within studies focused on beliefs about the *self* suggest that this is not a pivotal dimension for switching the good-news, bad-news effect on and off.

7.2 Addressing hedging

One inherent challenge in this literature studying belief updating in the presence of state-dependent stakes is the inherent hedging motive. Various approaches have been adopted to try to deal with it. Typically, the papers considering belief updating about an ego-related characteristic (e.g. IQ) rely on the (very reasonable) implicit assumption that hedging across the ego utility and monetary utility domains will be minimal.⁴⁸

Amongst the studies considering monetary state-dependent stakes, two different approaches to dealing with this challenge have been adopted. [Coutts \(2019\)](#) follows the method suggested in [Blanco et al. \(2010\)](#), which involves partitioning the world such that the participant will either be paid according to their belief *or* according to the state-dependent prize, never both. As discussed by [Blanco et al. \(2010\)](#) and Appendix A in [Coutts \(2019\)](#), this approach has some attractive theoretical properties. In contrast, in this paper we follow the method developed by [Offerman et al. \(2009\)](#) and [Kothiyal et al. \(2011\)](#). Both these two approaches are theoretically valid for alleviating the influence of hedging under the assumptions they make; both approaches have advantages and drawbacks. However, neither [Coutts \(2019\)](#) nor the current paper finds support for a larger responsiveness to good-news, which suggests that it is not the case that one of the two methods is allowing hedging to suppress the over-responsiveness to good-news.

In this paper, we chose to adopt the [Offerman et al. \(2009\)](#)-[Kothiyal et al. \(2011\)](#) method for the following reasons. Firstly, using QSR incentives facilitates a simple presentation of the incentives to subjects, potentially enhancing the understanding of the actual incentives faced. Secondly, under the assumptions discussed in the main text and appendices, the [Offerman et al. \(2009\)](#)-[Kothiyal et al. \(2011\)](#) method allows one to both correct for any hedging motive present, and to measure how large this hedging motive is in a particular incentive environment. Thirdly, as demonstrated by [Blanco et al. \(2010\)](#), hedging is generally not a major problem unless the hedging opportunity is transparent. To this end, we introduced the salience manipulation in the transparency of the hedging motive across our two ASYMMETRIC treatments. As advocated by [Blanco et al. \(2010\)](#), we see that the degree of hedging diminishes substantially in T3.SEPARATE when the hedging opportunity is not prominent. This suggests that even the uncorrected beliefs in T3.SEPARATE are a good approximation for the beliefs subjects hold. However, we also conduct all of our analysis using the corrected beliefs from each of the T2.COMBINED and T3.SEPARATE

⁴⁸For example, this assumes that subjects will not choose to artificially lower their reported belief about having a high IQ in order to hedge against actually having a low IQ. [Möbius et al. \(2014\)](#) use the crossover mechanism to ensure there is no hedging motive between the quizz *payment* and the reported belief, but otherwise rely on this implicit assumption of no hedging across domains. [Eil and Rao \(2011\)](#) use the QSR and state that they believe that subjects report their true beliefs since: (i) subjects are likely risk-neutral over the small belief elicitation incentives, (ii) they are principally told to report their beliefs, (iii) any tendency to hedge is likely to be less important for studying an asymmetry in belief updating, in comparison to studying the level of beliefs, since the hedging effects both the priors and the posteriors (and does so irrespective of whether good or bad news is observed). These three points apply also to the uncorrected beliefs in the current paper.

treatments. This provides two independent tests of the hypotheses we are testing, each of which is theoretically valid.

Even given these factors motivating our design choices, the possibility of hedging deserves attention when interpreting the results. However, when considering the full set of results we observe, there are several features of the results that suggest that hedging is not a driving factor behind the absence of a good-news, bad-news effect. The observed absence of any asymmetry in updating within either of the two *ASYMMETRIC* treatments for either the uncorrected or corrected beliefs, as well as the consistency in updating patterns observed across all three treatments is not easily explained by a combination of a ‘good-news, bad-news’ effect and hedging.

8 Conclusion

The objective of this paper was to study belief updating when an individual prefers one state of the world to another. In particular, the experiment was designed to test the *asymmetric updating hypothesis* in the domain of financial decision making, and contribute to the body of work that is constructing a descriptive understanding of how individuals form beliefs.

The main finding of the current paper is that we find no evidence for asymmetric updating in our experimental context with financial state-dependent stakes. Instead, we find that the updating behavior of the average individual is approximately Bayesian, irrespective of the presence or absence of financial stakes.

The current paper complements the existing work on this topic in several ways. Firstly, we consider belief updating from a wider range of prior beliefs for each individual. Secondly, we demonstrate how the exogeneity of the priors can be used to conduct empirical exercises that alleviate endogeneity concerns. Thirdly, our experimental design ensures that the distribution of realised signals observed is balanced in terms of the frequency of ‘good’ and ‘bad’ signals. This is useful as it removes the potential influence of the signal distribution that is documented by [Coutts \(2019\)](#). Fourth, our data allows us to conduct both *within subject* and *between subject* tests, as well as several robustness exercises in support of the main result.

The results described in this paper, and in this nascent literature as a whole, are instructive as there is a large class of economically important situations in which individuals form beliefs, preferring one state to another, ranging from capital markets to intertemporal portfolio choice problems to consumption savings problems. It is clearly important for economists to have an accurate understanding of how individuals form beliefs in these contexts. However, as described in the Discussion section above, the results observed in this literature are so far mixed. This paper contributes evidence towards revealing an accurate descriptive model of belief updating, showing an absence of the ‘good-news, bad-news’ effect in a context with financial stakes, no ego, and a

well defined and unambiguous information structure. However, to construct a more complete model of how we update our beliefs, more evidence is clearly needed.

References

- Akerlof, G. A. and W. T. Dickens (1982). The economic consequences of cognitive dissonance. *American Economic Review* 72(3), 307–319.
- Armantier, O. and N. Treich (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review* 62, 17–40.
- Barron, K. (2018). Belief updating: Does the 'good-news, bad-news' asymmetry extend to purely financial domains? *Working Paper*.
- Barron, K., S. Huck, and P. Jehiel (2019). Everyday econometricians: Selection neglect and overoptimism when learning from others. *Working Paper*.
- Benabou, R. and J. Tirole (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics* 117, 871–915.
- Bénabou, R. and J. Tirole (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives* 30(3), 141–164.
- Benjamin, D. (2019). *Handbook of Behavioral Economics - Foundations and Applications* 2, Chapter 2: Errors in probabilistic reasoning and judgment biases, pp. 69–186. North-Holland.
- Bernardo, A. E. and I. Welch (2001). On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy* 10(3), 301–330.
- Blanco, M., D. Engelmann, A. K. Koch, and H.-T. Normann (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics* 13(4), 412–438.
- Brunnermeier, M. and J. Parker (2005). Optimal expectations. *American Economic Review* 95(4), 1092–1118.
- Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2013). Overconfidence and social signalling. *Review of Economic Studies* 80(3), 949–983.
- Buser, T., L. Gerhards, and J. Van Der Weele (2018). Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty* 56(2), 165–192.
- Caplin, A. and J. Leahy (2001). Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of Economics* 116(1), 55–79.
- Chamley, C. (2004). *Rational Herds: Economic Models of Social Learning*. Cambridge University Press.

- Charness, G. and C. Dave (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior* 104, 1–23.
- Coutts, A. (2019). Good news and bad news are still news: experimental evidence on belief updating. *Experimental Economics*.
- Coutts, A., L. Gerhards, and Z. Murad (2019). No one to blame: Biased belief updating without attribution. *Working Paper*.
- Eil, D. and J. Rao (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics* 3(2), 114–138.
- Enke, B. (2019). What you see is all there is. *Working Paper*.
- Enke, B. and F. Zimmermann (2019). Correlation neglect in belief formation. *Review of Economic Studies* 86(1), 313–332.
- Ertac, S. (2011). Does self-relevance affect information processing? experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior & Organization* 80(3), 532–545.
- Esponda, I. and E. Vespa (2018). Endogenous sample selection: A laboratory study. *Quantitative Economics* 9(1), 183–216.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25(2), 383–417.
- Fama, E. F. (1976). Efficient capital markets: reply. *Journal of Finance* 31(1), 143–145.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171–178.
- Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton University Press.
- Garrett, N. and T. Sharot (2017). Optimistic update bias holds firm: Three tests of robustness following shah et al. *Consciousness and Cognition* 50, 12–22.
- Ghirardato, P., F. Maccheroni, and M. Marinacci (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory* 118(2), 133–173.
- Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18(2), 141–153.
- Gilboa, I. and D. Schmeidler (1993). Updating ambiguous beliefs. *Journal of economic theory* 59(1), 33–49.

- Gotthard-Real, A. (2017). Desirability and information processing: An experimental study. *Economics Letters* 152, 96–99.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association* 1:1, 114–125.
- Grether, D. M. (1978). Recent psychological studies of behavior under uncertainty. *American Economic Review* 68(2), 70–74.
- Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics* 95, 537–557.
- Grether, D. M. (1992). Testing bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization* 17(1), 31–57.
- Grossman, Z. and D. Owens (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization*.
- Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica* 75(2), 503–536.
- Hall, R. (1978). Stochastic implications of the life-cycle permanent income hypothesis. *Journal of Political Economy* 86, 971–987.
- Heger, S. and N. Papageorge (2018). We should totally open a restaurant: How optimism and overconfidence affect beliefs. *Journal of Economic Psychology*.
- Heidhues, P., B. Kőszegi, and P. Strack (2018). Unrealistic expectations and misguided learning. *Econometrica* 86(4), 1159–1214.
- Heifetz, A., C. Shannon, and Y. Spiegel (2007). The dynamic evolution of preferences. *Economic Theory* 32(2), 251.
- Hestermann, N. and Y. Le Yaouanq (2018). It's not my fault! self-confidence and experimentation. *Working Paper*.
- Holt, C. and A. Smith (2009). An update on bayesian updating. *Journal of Economic Behavior & Organization* 69(2), 125–134.
- Hossain, T. and R. Okui (2013). The binarized scoring rule. *Review of Economic Studies* 80(3), 984–1001.
- Jaffray, J.-Y. (1989). Linear utility theory for belief functions. *Operations Research Letters* 8(2), 107–112.

- Johnson, D. D. and J. H. Fowler (2011). The evolution of overconfidence. *Nature* 477(7364), 317–320.
- Kahneman, D. and A. Tversky (1973). On the psychology of prediction. *Psychological Review* 80(4), 237.
- Karni, E. (2009). A mechanism for eliciting probabilities. *Econometrica* 77(2), 603–606.
- Karni, E. and Z. Safra (1995). The impossibility of experimental elicitation of subjective probabilities. *Theory and Decision* 38(3), 313–320.
- Kothiyal, A., V. Spinu, and P. P. Wakker (2011). Comonotonic proper scoring rules to measure ambiguity and subjective beliefs. *Journal of Multi-Criteria Decision Analysis* 17, 101–113.
- Kuhnen, C. M. (2015). Asymmetric learning from financial information. *Journal of Finance* 70(5), 2029–2062.
- Kuzmanovic, B., A. Jefferson, and K. Vogeley (2015). Self-specific optimism bias in belief updating is associated with high trait optimism. *Journal of Behavioral Decision Making* 28(3), 281–293.
- Le Yaouanq, Y. and P. Schwardmann (2019). Learning about one’s self. *CESifo Working Paper*.
- Lewellen, J. and J. Shanken (2002). Learning, asset-pricing tests, and market efficiency. *Journal of Finance* 57(3), 1113–1145.
- Liang, Y. (2019). Learning from unknown information sources. *Working Paper*.
- Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal* 97, 666–684.
- Lucas, R. E. (1976). Econometric policy evaluation: A critique. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 1, pp. 19–46. North-Holland.
- Malmendier, U. and T. Taylor (2015). On the verges of overconfidence. *Journal of Economic Perspectives* 29(4), 3–8.
- Marks, J. and S. Baines (2017). Optimistic belief updating despite inclusion of positive events. *Learning and Motivation* 58, 88–101.
- Mayraz, G. (2013). The optimism bias: Experiment and model. *Working Paper*.
- Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics* 51, 247–257.
- Möbius, M., M. Niederle, P. Niehaus, and T. Rosenblat (2014). Managing self-confidence. *Working Paper*.

- Mossin, J. (1968). Optimal multiperiod portfolio policies. *Journal of Business* 41(2), 215–229.
- Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association* 55, 299–306.
- Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica* 29, 315–335.
- Ngangoue, K. (2018). Learning under ambiguity: An experiment on gradual information processing. *Working Paper*.
- Offerman, T., J. Sonnemans, G. van de Kuilen, and P. Wakker (2009). A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes. *Review of Economic Studies* 76(4), 1461–1489.
- Pires, C. P. (2002). A rule for updating ambiguous beliefs. *Theory and Decision* 53(2), 137–152.
- Prelec, D. (1998). The probability weighting function. *Econometrica* 66(3), 497–528.
- Rabin, M. and J. Schrag (1999). First impressions matter: A model of confirmatory bias. *Quarterly Journal of Economics* 114(1), 37–82.
- Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics* 51, 239–246.
- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* 57, 571–587.
- Schwardmann, P. and J. J. Van der Weele (2016). Deception and self-deception.
- Shah, P., A. Harris, U. Hahn, C. Catmur, and G. Bird (2016). A pessimistic view of optimistic belief updating. *Cognitive Psychology* 90, 71–127.
- Sharot, T., R. Kanai, D. Marston, C. W. Korn, G. Rees, and R. J. Dolan (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences* 109(42), 17058–17062.
- Sharot, T., C. W. Korn, and R. J. Dolan (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience* 14(11), 1475–1479.
- Trautmann, S. and G. van de Kuilen (2015). Belief elicitation: A horse race among truth serums. *Economic Journal* 125, 2116–2135.
- Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), 297–323.

APPENDICES

Appendix A: Robustness Checks

The empirical specification used in the main text of this paper assumes that updating follows the flexible parametric process described in Equation 1. This specification allows for a wide range of deviations from Bayes' rule, as discussed in Section 2. In this section we conduct several exercises to test for the robustness of the main results.

The first subsection examines whether the results from the main specification described in Equation 2 are robust to first differencing the dependent variable (i.e. this considers how new information moves the change in beliefs, imposing the assumption that $\delta = 1$). The second subsection extends the main empirical specification to allow for individual-specific updating parameters. The third subsection pools all the observations across the three treatments together, and then tests whether the average updating parameters differ across treatments, by interacting treatment group dummies with the regressors of the main specification described in Equation 2.

Table 4: First Difference Specification and Power Calculations

	T1: SYMMETRIC			T2: COMBINED			T3: SEPARATE		
	OLS	IV	DIFF	OLS	IV	DIFF	OLS	IV	DIFF
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
δ	0.90 (0.03)***	0.99 (0.03)		0.86 (0.04)***	0.99 (0.02)		0.93 (0.02)***	0.99 (0.02)	
γ_a	1.09 (0.11)	1.09 (0.11)	1.09 (0.11)	1.06 (0.12)	1.02 (0.11)	1.01 (0.11)	1.16 (0.11)	1.14 (0.11)	1.13 (0.10)
$\gamma_b - \gamma_a$	0.08 (0.08)	0.08 (0.08)	0.08 (0.08)	0.07 (0.10)	0.10 (0.09)	0.11 (0.09)	-0.03 (0.08)	-0.03 (0.08)	-0.02 (0.08)
$p(\gamma_a = \gamma_b)$	0.32	0.31	0.32	0.48	0.25	0.25	0.73	0.74	0.75
MDE ($\kappa = 0.8$)	0.24	0.22	0.22	0.27	0.25	0.26	0.24	0.22	0.22
R^2	0.73		0.31	0.74		0.21	0.84		0.34
1st Stage F		84.04			107.01			95.45	
N	1,075	1,075	1,075	1,285	1,285	1,285	1,140	1,140	1,140

(i) Standard errors in parentheses (clustered at the individual level)

(ii) T-tests of $H_0: \delta = 1; \gamma_a = 1; \gamma_b - \gamma_a = 0$ indicated by * = 10%, ** = 5%, *** = 1%

(iii) MDE reports the minimum detectable effect size for a power of κ .

Robustness Check 1: First-Differences Specification and Power Calculation

This section of the robustness checks serves two purposes. The first purpose is to check for the robustness of the results from the core empirical specification to the use of a first differences specification (DIFF), which essentially involves imposing the assumption that $\delta = 1$. The second purpose of this section is to report the size of the minimum detectable effect (MDE) from power calculations for both our main OLS and IV empirical specification, and the DIFF specification.

One of the challenges in carrying out a statistical analysis of belief updating behavior is that an individual's current posterior belief necessarily depends upon her prior belief, which in turn is the result of updating in response to past information. Therefore, when estimating a parametric belief updating function, one concern is that the individual's prior belief is correlated with unobservables. In the main text, we devoted substantial space to discussing how the experiment was designed explicitly to address this concern by generating a completely exogenous information set, facilitating a natural instrumental variables (IV) approach to estimation. The first differences specification results presented here serve to further complement the IV analysis, since the DIFF specification avoids the potential endogeneity issue by removing the lagged belief from the set of dependent variables in the regression.

In columns (#a) and (#b), Table 4 repeats the OLS and IV results from Table 3 for the corrected beliefs, with one minor change to the core specification in Equation 2. Here we report, instead, the results for the equivalent specification:

$$\tilde{\pi}_{i,j,t+1} = \delta \tilde{\pi}_{i,j,t} + \gamma_a \hat{q} - (\gamma_b - \gamma_a) \tilde{q} \cdot 1(s_{i,j,t+1} = b) + \epsilon_{i,j,t+1} \quad (3)$$

where $\tilde{\pi}_{i,j,t} = \text{logit}(\pi_{i,j,t})$ and $\hat{q} = \log(\frac{q}{1-q}) \cdot [1(s_{i,j,t} = a) - 1(s_{i,j,t} = b)]$; while as above, $\tilde{q} = \log(\frac{q}{1-q})$; j refers to a round of decisions; t counts the decision numbers within a round, and the errors ϵ_{ijt+1} are clustered at the individual (i) level. The difference $\gamma_b - \gamma_a$ denotes a single parameter estimated in the regression, but is denoted as the difference between γ_b and γ_a as this is the natural way to think about this parameter in the context of the discussion above (i.e. the difference between how subjects update in response to 'bad news' and 'good news').

The reason for the rearrangement of the equation is that, while it is equivalent⁴⁹ to the specification in Equation 2, it displays the test of the difference between γ_a and γ_b more clearly (i.e. the

⁴⁹Notice that the regression coefficients and standard errors on δ and γ_a are the same in Tables 3 and 4 (where we are only considering the corrected beliefs). Furthermore, we can see the equivalence from the following simple rearrangement:

$$\begin{aligned} \tilde{\pi}_{i,j,t+1} &= \delta \tilde{\pi}_{i,j,t} + \gamma_a \tilde{q} \cdot 1(s_{i,j,t+1} = a) - \gamma_b \tilde{q} \cdot 1(s_{i,j,t+1} = b) \\ &= \delta \tilde{\pi}_{i,j,t} + \gamma_a \tilde{q} \cdot 1(s_{i,j,t+1} = a) - \gamma_a \tilde{q} \cdot 1(s_{i,j,t+1} = b) + \gamma_a \tilde{q} \cdot 1(s_{i,j,t+1} = b) - \gamma_b \tilde{q} \cdot 1(s_{i,j,t+1} = b) \\ &= \delta \tilde{\pi}_{i,j,t} + \gamma_a \tilde{q} \cdot [1(s_{i,j,t+1} = a) - 1(s_{i,j,t+1} = b)] + [\gamma_a - \gamma_b] \cdot \tilde{q} \cdot 1(s_{i,j,t+1} = b) \end{aligned}$$

test of the *asymmetric updating hypothesis*), and thereby also facilitates calculating the MDE. In Table 3, we have presented the MDE for a power of $\kappa = 0.8$.

Columns (#c) report the results for the first difference specification, which imposes the restriction that $\delta = 1$:

$$\Delta \tilde{\pi}_{i,j,t+1} = \gamma_a \hat{q} - (\gamma_b - \gamma_a) \tilde{q} \cdot 1(s_{i,j,t+1} = b) + \epsilon_{i,j,t+1} \quad (4)$$

where $\Delta \tilde{\pi}_{i,j,t+1} = \text{logit}(\pi_{i,j,t+1}) - \text{logit}(\pi_{i,j,t})$ and $\hat{q} = \log(\frac{q}{1-q}) \cdot [1(s_{i,j,t} = a) - 1(s_{i,j,t} = b)]$; j refers to a round of decisions; t counts the decision numbers within a round, and the errors ϵ_{ijt+1} are clustered at the individual (i) level.

The results indicate that the $\gamma_b - \gamma_a$ parameter is robust to the different empirical specifications adopted, and also doesn't vary substantially across treatment groups. In all treatment groups, and for each of the empirical specifications considered, we cannot reject the null hypothesis that this parameter is equal to zero, which implies that we do not find support for the *asymmetric updating hypothesis*. Furthermore, we calculate the MDE for each specification, considering a significance level of $\alpha = 0.05$ and a power of $\kappa = 0.8$. Under these assumptions, the MDE for the difference between the γ_b and γ_a parameters in each of the regressions considered in isolation ranges between 0.22 and 0.27. As a result, we cannot conclusively reject the possibility that there exists a small asymmetry in updating; however none of our results provide any support for this conclusion.

Robustness Check 2: Allowing for Individual-Specific Updating Parameters

As discussed above, one reason we might think that endogeneity of the lagged belief could lead to biased estimates is if there is heterogeneity in individual updating behavior and this leads to a correlation between the unobserved error term and the lagged belief variable amongst the regressors. We have tried to address this issue above using, firstly, an instrumental variable approach, and secondly, a first differences empirical specification. However, since the data were collected in the form of a panel of belief updates for each individual, the data lends itself to controlling for individual-specific behavior through exploiting the panel. A typical fixed effects model is not appropriate here, as it is not the *level* of the regression that shifts from individual to individual. However, we can include individual-specific updating parameters to control for the *slope* to shift at the individual level. This allows us to extract the individual heterogeneity in how responsive individuals are to their prior belief, and to new information in general, and reduce the possible bias in the main parameter of interest, the average difference in responsiveness to 'bad news' and 'good news': $\gamma_b - \gamma_a$. With this in mind, our third robustness check involves estimating the following empirical specification:

$$\tilde{\pi}_{i,j,t+1} = \delta_i \tilde{\pi}_{i,j,t} + \gamma_i \hat{q} - (\gamma_b - \gamma_a) \tilde{q} \cdot 1(s_{i,j,t+1} = b) + \epsilon_{i,j,t+1} \quad (5)$$

where δ_i and γ_i are estimated at the individual level, and the remaining parameters and variables are defined as above. The results from this exercise using the corrected beliefs are reported in Table 5.

These results are very consistent with the estimates from the core specification, as well as from the DIFF specification in Robustness Check 2. In summary, all the empirical estimates provide support for the same underlying story that the data collected in this experiment provide no support for the *asymmetric updating hypothesis* in this context.

Table 5: Allowing for Individual-Specific Updating Parameters.

	T1 SYMMETRIC	T2 COMBINED	T3 SEPARATE	T2+T3
	(1)	(2)	(3)	(4)
$\gamma_b - \gamma_a$	0.09 (0.10)	0.14 (0.10)	-0.04 (0.09)	0.06 (0.07)
$p(\gamma_a = \gamma_b)$	0.35	0.18	0.67	0.41
MDE ($\kappa = 0.8$)	0.28	0.29	0.26	0.20
MDE ($\kappa = 0.9$)	0.33	0.33	0.30	0.23
N	1,075	1,285	1,140	2,425
R^2	0.80	0.84	0.89	0.86

(i) Standard errors in parentheses

(ii) T-tests of H_0 : Coefficient = 0 reported: * = 10%, ** = 5%, *** = 1%

(iii) MDE reports the minimum detectable effect size for a power of κ .

Robustness Check 3: Between Treatments Comparison of Updating Parameters

This section tests whether the belief updating parameters in our core specification are significantly different *between* the three treatment groups. This is done by pooling together the three treatment groups and estimating Equation 2, but with the inclusion of treatment dummies interacted with the updating coefficients. This provides us with a test of whether the parameters differ between either of the two ASYMMETRIC treatments and SYMMETRIC.

More specifically, this involves estimating the following equation:

$$\begin{aligned}\tilde{\pi}_{i,j,t+1} = & \delta \tilde{\pi}_{i,j,t} + \gamma_a \tilde{q} \cdot 1(s_{i,j,t+1} = a) - \gamma_b \tilde{q} \cdot 1(s_{i,j,t+1} = b) + \\ & \sum_{k=2}^3 [\delta^k \tilde{\pi}_{i,j,t} \cdot T_{i,j,t}^k + \gamma_a^k \tilde{q} \cdot 1(s_{i,j,t+1} = a) \cdot T_{i,j,t}^k - \gamma_b^k \tilde{q} \cdot 1(s_{i,j,t+1} = b) \cdot T_{i,j,t}^k] + \epsilon_{i,j,t+1}\end{aligned}$$

where $T_{i,j,t}^k$ is an indicator variable for treatment k [i.e. $T_{i,j,t}^k = 1(T_{i,j,t} = k)$], with $T_{i,j,t}$ a treatment variable taking the values $\{1, 2, 3\}$ corresponding to the three treatment groups. The coefficients δ, γ_a , and γ_b reflect the baseline parameters without the influence of state-contingent stakes and the parameters δ^k, γ_a^k and γ_b^k estimate the movement from these parameters for each of the two state-contingent stake treatments, $k \in \{2, 3\}$.

The results from this exercise are presented in Table 6. The results show that, for the average individual, there are no systematic differences in the updating parameters across treatment groups. This implies that the differences in exogenous state-contingent incentives do not exert a strong influence on how individuals update their beliefs in the different treatments.

Table 6: Testing for Differences in Average Updating Behavior *between* Treatment Groups.

	Belief 1 (1)	Belief 2 (2)	Belief 3 (3)	Belief 4 (4)	Belief 5 (5)	Pooled (6)	Full Sample (7)
<u>Priors</u>							
δ	0.93 (0.05)	0.89 (0.04)	1.03 (0.05)	1.00 (0.06)	1.02 (0.06)	0.99 (0.03)	1.00 (0.03)
$\delta * T2$	-0.02 (0.07)	0.02 (0.07)	-0.02 (0.07)	0.01 (0.07)	0.01 (0.07)	-0.00 (0.04)	-0.04 (0.03)
$\delta * T3$	0.02 (0.08)	0.06 (0.06)	-0.10 (0.06)	0.08 (0.07)	-0.04 (0.09)	0.00 (0.04)	-0.00 (0.03)
<u>Signal: Blue ($s = a$)</u>							
γ_a	0.81 (0.14)	1.02 (0.18)	1.03 (0.18)	1.30 (0.21)	1.28 (0.21)	1.09 (0.11)	0.84 (0.09)
$\gamma_a * T2$	0.09 (0.22)	-0.07 (0.26)	-0.15 (0.22)	-0.08 (0.34)	-0.16 (0.29)	-0.07 (0.15)	-0.01 (0.13)
$\gamma_a * T3$	0.12 (0.21)	-0.05 (0.22)	-0.10 (0.20)	0.07 (0.28)	0.21 (0.31)	0.05 (0.15)	0.04 (0.13)
<u>Signal: Red ($s = b$)</u>							
γ_b	0.69 (0.11)	0.79 (0.16)	1.18 (0.19)	1.40 (0.21)	1.80 (0.23)	1.16 (0.11)	0.78 (0.09)
$\gamma_b * T2$	0.21 (0.20)	0.06 (0.22)	-0.33 (0.25)	-0.21 (0.27)	0.14 (0.40)	-0.04 (0.17)	0.11 (0.15)
$\gamma_b * T3$	0.30 (0.18)	0.04 (0.20)	-0.15 (0.21)	-0.37 (0.25)	-0.06 (0.33)	-0.05 (0.15)	0.10 (0.12)
Observations	700	700	700	700	700	3,500	5,550
Kleibergen-Paap F	46.61	47.46	56.63	60.45	64.03	93.17	53.88

(i) Robust standard errors in parentheses (clustered at the individual level).

(ii) Estimates use the corrected beliefs and are instrumented using the correct lagged Bayesian posterior.

(iii) All of the non-interacted coefficients are significantly different from 0 at the 1% level. Only one of the forty-two interaction coefficients are significantly different from zero at the 10% level. This is the $\gamma_b * T3$ coefficient in the Belief 1 column, which is significant at the 10%, but not the 5% level.

Appendix B.1: Core Properties of Bayes' Rule

Möbius et al. (2014) argue that the core structure of Bayesian updating is captured by the following three properties:

1. *invariance*, whereby the difference in logit beliefs between t and $t + 1$ depends only on the history of signals, H_{t+1} , and the initial prior, p_0 (i.e. on the agent's information set). An updating process is *invariant* if we can find a function g_t such that:

$$\text{logit}(\pi_{t+1}) - \text{logit}(\pi_t) = g_t(s_{t+1}, s_t, \dots, s_1; p_0)$$

If an individual displayed base rate neglect or confirmatory bias, this would constitute a violation of invariance (i.e. in the context of the model outlined in the main text, this assumption stipulates that $\delta = 1$).

2. π_t is a *sufficient statistic* for all information received at time t or earlier, such that the change in logit beliefs depends only on the new information in time $t + 1$: $\text{logit}(\pi_{t+1}) - \text{logit}(\pi_t) = g_t(s_{t+1})$
3. *stability* of the updating process over time. This property is satisfied if $g_t = g$ for all t .

Under the assumption that these properties are satisfied, the authors note that the class of updating processes that remain can be fully described by the two parameter function, $g(s_t)$, where:

$$g(s_t) = \log\left(\frac{q}{1-q}\right) \cdot 1(s_{t+1} = a) - \log\left(\frac{q}{1-q}\right) \cdot 1(s_{t+1} = b)$$

This serves to motivate the model described in Equation 1.

Appendix B.2.1: The QSR and a Non-EU 'Truth Serum'

In this section, we discuss how beliefs reported under the QSR might be distorted, and how we address this challenge. Consider the binary event, denoted by E_ω , where $\omega \in \{A, B\}$. Therefore, E_A refers to the event that state $\omega = A$ is realized. The object that we would like to elicit is the participant's belief, $\pi_t = P(E_A) = P(\omega = A)$, regarding the likelihood that state $\omega = A$ is the correct state at time t . However, the object that we will observe is the participant's reported belief, r_t , at each point in time under the incentives prescribed by the quadratic scoring rule. The *Quadratic Scoring Rule* at time t is defined by:

$$S_A(r_t) = 1 - (1 - r_t)^2 \quad (6)$$

$$S_B(r_t) = 1 - r_t^2 \quad (7)$$

where r_t is the reported probability of event E_A occurring; $S_A(r_t)$ is the payment if the state $\omega = A$ is realized; $S_B(r_t)$ is the payment if the state $\omega = B$ is realized. Therefore, the QSR essentially involves a single choice from a list of binary prospects, $(1 - (1 - r_t)^2)_{E_A}(1 - r_t^2)$. The QSR is a ‘proper’ scoring rule since, if the agent is a risk neutral EU maximizer then she is incentivized to truthfully reveal her belief, π_t :

$$\pi_t = \arg \max_{r_t \in [0,1]} \pi_t S_A(r_t) + (1 - \pi_t) S_B(r_t)$$

However, the QSR is no longer incentive compatible once we allow for (i) *risk aversion / loving* and (ii) participants who have exogenous stakes in the state of the world. The reasons for this are the following. Firstly, it has been well documented theoretically that, if the participant is *risk averse*, then the QSR leads to reporting of beliefs, r_t , that are distorted towards 0.5, away from her true belief, π_t , when the participant has no exogenous stakes in the realized state.⁵⁰ This distortion has been observed in experimental data (Offerman et al., 2009; Armantier and Treich, 2013). Secondly, in our experiment, we will also be interested in eliciting beliefs when participants have an exogenous stake associated with one of the two states. More precisely, we will be interested in recovering the participant’s true belief when she receives an exogenous payment, x , if state $\omega = A$ is realized. This payment, x , is in addition to the payment she receives from the QSR. In other words, she will choose from a menu of binary prospects of the form: $(x + 1 - (1 - r_t)^2)_{E_A}(1 - r_t^2)$. In the context of state-dependant stakes, a risk averse EU maximizer⁵¹ will face two distortionary motives in reporting her belief: (i) she will face the motive to distort her belief towards 0.5 as discussed above; and (ii) in addition, there is a hedging motive, which will compel a risk averse individual to *lower* her reported belief, r_t , towards zero as x increases.

If the participants in our experiment are *risk neutral expected utility* maximizers, the reported beliefs, r_t , that we elicit under the QSR will coincide with their true beliefs, π_t . However, in order to allow for choice behaviors consistent with a wider range of decision models, we will measure the size of the distortionary influence of the elicitation incentives at an individual level and correct the

⁵⁰i.e. if $\pi_t > 0.5$ then $\pi_t > r_t > 0.5$, and if $\pi_t < 0.5$ then $\pi_t < r_t < 0.5$ for a risk averse individual reporting her beliefs under QSR incentives.

⁵¹A participant who is a risk averse EU maximizer will choose her reported belief r_t by solving the following maximization problem:

$$\max_{r_t \in [0,1]} \pi_t U(x + 1 - (1 - r_t)^2) + (1 - \pi_t) U(1 - r_t^2)$$

beliefs accordingly. This approach is valid under the weak assumption that individuals evaluate binary prospects according to the *biseparable preferences*⁵² model and are *probabilistically sophisticated*.⁵³ This restriction on behavior is very weak and includes individuals who behave according to EU with any risk preferences as well as the majority of commonly used NEU models.⁵⁴

A Non-EU ‘Truth Serum’

The discussion above has highlighted how beliefs might be distorted under QSR incentives. The Offerman et al. (2009) approach proposes correcting the reported beliefs for the risk aversion caused by the curvature of the utility function or by non-linear probability weighting. This approach involves eliciting participants’ reported belief parameter, r , for a set of risky events where they know the objective probability, p (*known probability*). This is done under precisely the same QSR incentive environment in which we elicit the participants’ subjective beliefs, π , regarding the events of interest (where they don’t know the objective probability: *unknown probability*). If a subject’s reported beliefs, r , differ from the known objective probabilities, p , this indicates that the subject is distorting her beliefs due to the incentive environment (e.g. due to risk aversion). The objective of the correction mechanism is therefore to construct a map, R , from the objective beliefs, $p \in [0, 1]$, to the reported beliefs, r , for each individual under the relevant incentive environment.

Offerman et al. (2009) show that under the assumption that individuals evaluate prospects in a way that is consistent with the weak assumptions of the *biseparable preferences* model, then in the scenario where there are no state-contingent stakes (i.e. $x = 0$), individuals evaluate the QSR menu of prospects $(1 - (1 - r_t)^2)_{E_A}(1 - r_t^2)$ according to $w(P(E_A))U(1 - (1 - r_t)^2) + (1 - w(P(E_A)))U(1 - r_t^2)$ for $r_t \geq 0.5$ and therefore the inverse of the map from objective probabilities

⁵²The *biseparable preference* model holds if the preference ordering, \succsim , over prospects of the form, $y_E z$, can be represented by:

$$y_E z \rightarrow W(E)U(y) + (1 - W(E))U(z)$$

where U is a real-valued function unique up to level and unit; and W is a unique weighting function, satisfying $W(\emptyset) = 0$, $W(S) = 1$ and $W(E) \leq W(F)$ if $E \subseteq F$. S is the set of all states and events are subsets of the full set of states: i.e. $E, F \subseteq S$. In this paper, we only consider two-state prospects, where the state-space is partitioned into two parts by an event, E and its complement E^c . Making the further assumption that the decision maker is *probabilistically sophisticated* gives the following refinement:

$$y_E z \rightarrow w(P(E))U(y) + (1 - w(P(E)))U(z)$$

⁵³*Probabilistic sophistication* is the assumption that we can model that individual’s preferences over prospects as if the individual’s beliefs over states can be summarized by a probability measure, P . In other words, probabilistic sophistication implies that we can model the individual’s belief regarding the likelihood of an event E as being completely summarized by a single probability judgment, $P(E_A)$.

⁵⁴Amongst the models subsumed within the biseparable preferences model are EU, Choquet expected utility (Schmeidler, 1989), maxmin expected utility (Gilboa and Schmeidler, 1989), prospect theory (Tversky and Kahneman, 1992), and α -maxmin expected utility (Ghirardato et al., 2004). See Offerman et al. (2009) for a discussion.

to reported probabilities, R , is given by:

$$p = R^{-1}(r) = w^{-1} \left(\frac{r}{r + (1-r) \frac{U'(1-(1-r)^2)}{U'(1-r^2)}} \right) \quad (8)$$

In Appendix B.2.2, we provide a derivation for this equation, as well as augmenting the Offerman et al. (2009) approach to allow for the scenario where there are state-contingent stakes (i.e. $x \neq 0$). This extension to Offerman et al. (2009) represents a special case of the more general treatment of correction methods for binary proper scoring rules considered by Kothiyal et al. (2011). In our empirical analysis, we discuss how we use Equation 8 to recover the function, R , for each individual and thereby recover their beliefs, π_t , from their reported beliefs, r_t .

Appendix B.2.2: Augmenting the Offerman et al. (2009) ‘Truth Serum’ Approach to Include Stakes

The previous section discussed the central ideas motivating the Offerman et al. (2009) approach for correcting for hedging in cases where there are no state-dependent stakes (i.e. $x = 0$). In projects studying the *asymmetric updating hypothesis*, allowing for state-dependent stakes (i.e. $x \neq 0$) is of fundamental importance. Therefore, in this section, we consider an extension to the Offerman et al. (2009) approach to correcting for hedging. The extension we consider is tailored specifically to our experimental setting, however it is a special case of the more general set of correction techniques studied by Kothiyal et al. (2011).⁵⁵

In the case where $x \neq 0$, the text above discussed how participants who face the quadratic scoring rule incentives, along with the non-zero state-contingent bonus x , essentially face a choice from a menu of lotteries denoted by $(x + 1 - (1 - r_t)^2)_{EA}(1 - r_t^2)$. An individual who satisfies the *biseparable preferences* model and is probabilistically sophisticated will evaluate this prospect using the following Equations:⁵⁶

For $x \geq 1$ or $r_t \geq 0.5$:

$$w(P(E_A))U(x + 1 - (1 - r_t)^2) + (1 - w(P(E_A)))U(1 - r_t^2) \quad (9)$$

⁵⁵Kothiyal et al. (2011) extend the basic idea used by techniques aiming to correct elicited beliefs for reporting bias (e.g. hedging) to apply to the set of all binary proper scoring rules, and cover the full domain of beliefs. In conjunction with Offerman et al. (2009), this paper therefore offers a useful set of tools for accessing subjects’ true beliefs in situations where they may have reason to distort their reports.

⁵⁶For expositional simplicity, we don’t consider $x \in (0, 1)$. The discussion below is easily extended to these cases, but they are irrelevant for the purposes of this paper. This case is slightly different due to the fact that the probability weights on events or states may depend on their ordinal ranking according to preferences in this model.

and similarly,

For $x = 0$ & $r_t < 0.5$:

$$(1 - w(P(E_A^c)))U(x + 1 - (1 - r_t)^2) + w(P(E_A^c))U(1 - r_t^2) \quad (10)$$

The reason for the two separate conditions is due to the way in which many NEU models, subsumed within *biseparable preferences* model, allow the probability weighting function, $w(\cdot)$, over events to be influenced by the ordinal ranking over the associated outcomes, from best to worst⁵⁷. Since the case where $x = 0$ is discussed extensively in [Offerman et al. \(2009\)](#), we will focus on the case where $x \geq 1$ in the discussion that follows. This case only requires a very minor adjustment to their discussion. The key results are the following (adjusted to include the influence of x):

Result 1: Under NEU with **known probabilities**, p , the optimal reported probability, $r = R_x(p)$ satisfies:

$$\text{If } x \geq 1, \text{ then } p = R_x^{-1}(r_t) = w^{-1} \left(\frac{r_t}{r_t + (1 - r_t) \frac{U'(x+1-(1-r_t)^2)}{U'(1-r_t^2)}} \right) \quad (11)$$

Result 2: Under NEU with **unknown probabilities**, the optimal reported probability, r , satisfies:

$$\text{If } x \geq 1, \text{ then } P(E) = w^{-1} \left(\frac{r_t}{r_t + (1 - r_t) \frac{U'(x+1-(1-r_t)^2)}{U'(1-r_t^2)}} \right) \quad (12)$$

This motivates the simple strategy for recovering the agent's subjective beliefs from her reported beliefs under the specific incentive environment that she faces. Since the RHS of (11) and (12) agree, we have:

$$P(E) = R_x^{-1}(r) \quad (13)$$

which implies that if we can recover the function R_x^{-1} then we can map the reported beliefs to the participant's subjective beliefs. Equation 11 shows that we can recover this R_x^{-1} function in the same way here, with the bonus payment of x , as in the case where $x = 0$ considered in [Offerman et al. \(2009\)](#). Essentially, we provide the participant with prospects over known probabilities, p , and ask them for their belief regarding the likelihood that one state will be realized. In order to ensure the incentives to distort one's reported beliefs are kept constant, we do this exercise under precisely the same incentive environment as in the main belief updating task. By eliciting these

⁵⁷When $x = 0$ and $r_t < 0.5$, then $1 - r_t^2 > 1 - (1 - r_t)^2$ (i.e. in this case, E_A^c becomes the preferred event, rather than E_A , and therefore the probability weighting function is reversed).

reported beliefs associated with known probabilities spanning the whole unit interval, we can use these (p, r_t) pairs to estimate $R_x(p)$ for each individual for the relevant incentive environment created by the belief elicitation. Having estimated $R_x(p)$, we can calculate its inverse, $R_x^{-1}(r)$.

We can take any beliefs reported by the participant under the same belief elicitation incentives and then use this estimated $R_x^{-1}(r)$ to recover her true beliefs. In particular, we can use this estimated function to recover her true beliefs from her reported beliefs in the belief updating task that is the focus of this paper. Essentially, we are using this procedure to remove any misreporting effect that the belief elicitation incentive environment may have. It allows us to correct for the possibility that individuals may hold some belief $P(E)$ or π , but instead report a different belief, r .

If the incentive environment does not cause the participant to report a belief different from her true belief, then this procedure is unnecessary, but applying the procedure to her reported beliefs will not have any effect. In this case, the corrected beliefs will be the same as the reported beliefs.

Appendix B.2.3: Calibration of the Belief Correction Procedure: Theory

It is clear from the discussion in the main text and Equation 8, that we could recover $R(\cdot)$ *non-parametrically* for each individual if we were to collect a large number of (p, r) pairs from participants, such that the interval between the known probabilities, p , is sufficiently small. However, since it is not practical here to elicit such a large number of observations from each participant, we instead impose a parametric structure similar to the one used by [Offerman et al. \(2009\)](#).

For the utility function, $U(\cdot)$, we use the *constant relative risk aversion (CRRA)*⁵⁸ functional form:

$$U(x) = \begin{cases} x^\rho & \text{if } \rho > 0 \\ \ln x & \text{if } \rho = 0 \\ -x^\rho & \text{if } \rho < 0 \end{cases} \quad (14)$$

For the probability weighting function, $w(\cdot)$, we adopt Prelec's (1998), one-parameter family⁵⁹:

⁵⁸It is worth noting that the name "constant relative risk aversion (CRRA)" is not appropriate here as risk attitudes are also captured in the probability weighting function.

⁵⁹This is a special case of Prelec's two-parameter family of weighting functions:

$$w(p) = \exp[-\beta(-\ln(p))^\alpha]$$

For the purposes of the current context, the two-parameter family is not practically suitable due to the limited data we use at the individual level. This functional form permits the standard inverse-S shaped probability weighting

$$w(p) = \exp[-(-\ln(p))^\alpha] \quad (15)$$

Substituting these parametric functional form specifications into Equation 8 for the case where $x = 0$ gives:

$$p = R^{-1}(r_t) = \exp \left(- \left[-\ln \left(\frac{r_t(2r_t - r_t^2)^{1-\rho}}{r_t(2r_t - r_t^2)^{1-\rho} + (1-r_t)(1-r_t^2)^{1-\rho}} \right) \right]^{\frac{1}{\alpha}} \right) \quad (16)$$

For the the case where $x \neq 0$, substituting these parametric functional form specifications described in Equations 14 and 15 into Equation 11 gives:

$$p = R_x^{-1}(r_t) = \exp \left(- \left[-\ln \left(\frac{r_t(1 - (1-r_t)^2 + x)^{1-\rho}}{r_t(1 - (1-r_t)^2 + x)^{1-\rho} + (1-r_t)(1-r_t^2)^{1-\rho}} \right) \right]^{\frac{1}{\alpha}} \right) \quad (17)$$

We therefore use this adapted specification for our correction mechanism for the COMBINED and SEPARATE treatment groups.

In our core analysis, for our individual level reported belief corrections, we will make the simplifying assumption that $\alpha = 1$, such that risk aversion is captured only through the curvature of the utility function and not through the probability weighting function. The results are similar when we use Prelec's one parameter weighting function. Furthermore, it is substantially easier to interpret the risk aversion parameter estimates when we estimate ρ alone, due to the strong relationship between the ρ and α estimates .

We therefore estimate the following model, for each participant, in order to acquire a numerical estimate for the inverse of this function, $R(\cdot)$:

$$\text{logit}[R(j/20)] = \text{logit}[h(j/20, \alpha, \rho)] + u_j \quad (18)$$

where $R(j/20)$ is the probability reported by the individual that corresponds to true known probability, $p = \frac{j}{20}$ where $1 \leq j \leq 19$ ⁶⁰. As discussed above, α is the parameter of the probability weighting function; ρ gives the curvature of the utility function. The function $h(\cdot)$ is the inverse of R^{-1} . We estimate this function, $h(\cdot)$ numerically at each step⁶¹ within the maximum likelihood estimation. The error terms, u_j , are independently and identically distributed across participants

function that has been found to be consistent with the majority of the existing empirical evidence. When $\beta = 1$ in the one-parameter family, the α parameter captures the degree of curvature of the inverse-S shape but the point at which $w(p)$ intersects the 45 degree line is predetermined. Adding the second parameter, β , extends the one-parameter specification by allowing this fixed point to vary.

⁶⁰In other words, for known probabilities, p , between 0.05 and 0.95 at intervals of 0.05.

⁶¹i.e. given the current parameter guesses.

and choices and are drawn from a normal distribution. Essentially, here we are using each participant's 20 (r, p) pairs in order to estimate an R function that reflects the distortion in her reported beliefs due to the particular quadratic scoring rule incentive structure that she is subject to. Notice, that this structure varies across treatments as x varies and therefore the same subject would require a different adjustment curve if she were reassigned to a different treatment.

Using these estimates at an individual level allows us to recover the participants' true subjective beliefs, π_t , from the first stage of the experiment in which they report their beliefs, r_t , regarding the likelihood of $\omega = A$ being the true state. In Figure 6 in the main text, we graph the individual level correction curve estimates for two individuals in each treatment group. It is clear from these examples, firstly, that individuals in the sample are distorting their reported beliefs substantially relative to the known probabilities, and secondly, that the estimated correction curves are sufficiently flexible to fit different types of belief distortion behavior reasonably well. Furthermore, importantly, the graph in the top-left panel of the figure shows that, when an individual accurately reports her beliefs, then the correction mechanism has no harmful effect.

At the aggregate level, for each treatment group, $T \in \{1, 2, 3\}$, we estimate:

$$\text{logit}[R_i(j/20)] = \text{logit}[h_i(j/20, \alpha, \rho)] + u_{i,j} \quad (19)$$

where j indexes the 20 reported probabilities of individual i . This specification allows us to examine the distortion caused by the incentive environment to the average individual in each of the three treatment groups.

Appendix B.2.4: Calibration of the Belief Correction Procedure: Estimation

The belief correction procedure that we adopt involves assuming a flexible parametric form for the participants' utility and probability weighting functions in order to estimate the R function discussed in Equation 8 above. We estimate this function for each individual separately in order to correct the reported beliefs at the individual level. In addition, we estimate this function at the aggregate level for each of the treatment groups in order to obtain a measure of the average distortion of the incentive environment faced in each of the treatment groups. A detailed discussion of the mechanics of the Belief Correction Procedure we use is provided in Appendix B.2.3 above. Essentially, we are simply fitting a curve through each subject's belief elicitation incentive distortion.

Figure 5 displays the average correction curves for each of the treatment groups, fitting a single curve to the reported belief data observed across all subjects in the relevant treatment group.

Comparing the three subgraphs, we see that the average individual distorts the beliefs she reports in a way that is consistent with what risk aversion under EU would predict, with the inverse-S shape distortion in the T1.SYMMETRIC stakes treatment and the strong distortion downwards (away from the more desirable state) in both of the ASYMMETRIC stakes treatments. Furthermore, we see that the different ways of framing the same incentives in the two ASYMMETRIC treatments has a clear influence on behavior. In T2.COMBINED, the participants hedge far more when choosing their reported beliefs in comparison to those in the T3.SEPARATE group. This is in spite of the fact that the incentives are identical in these two treatments. This indicates that the reported beliefs in the T3.SEPARATE treatment are closer to the participants' true beliefs and motivates this presentation of incentives as preferable for future work that calls for the elicitation of beliefs when there are exogenous state-contingent payments.

At the individual level, there is a large degree of heterogeneity in the degree to which individuals distorted their reported belief away from their actual belief, given the incentive environment. Figure 6 displays the correction curves estimated for two individuals from each treatment group. It is clear from this figure that some individuals responded very strongly to the incentive environment in which their belief was elicited, while others reported their belief more accurately. The belief correction procedure is therefore very helpful for recovering the true beliefs of participants who responded strongly to the incentive environment. In cases where the individual simply reported their belief accurately, the corrected beliefs and the reported beliefs are exactly the same.

Figure 5: Average Correction Functions across Treatments.

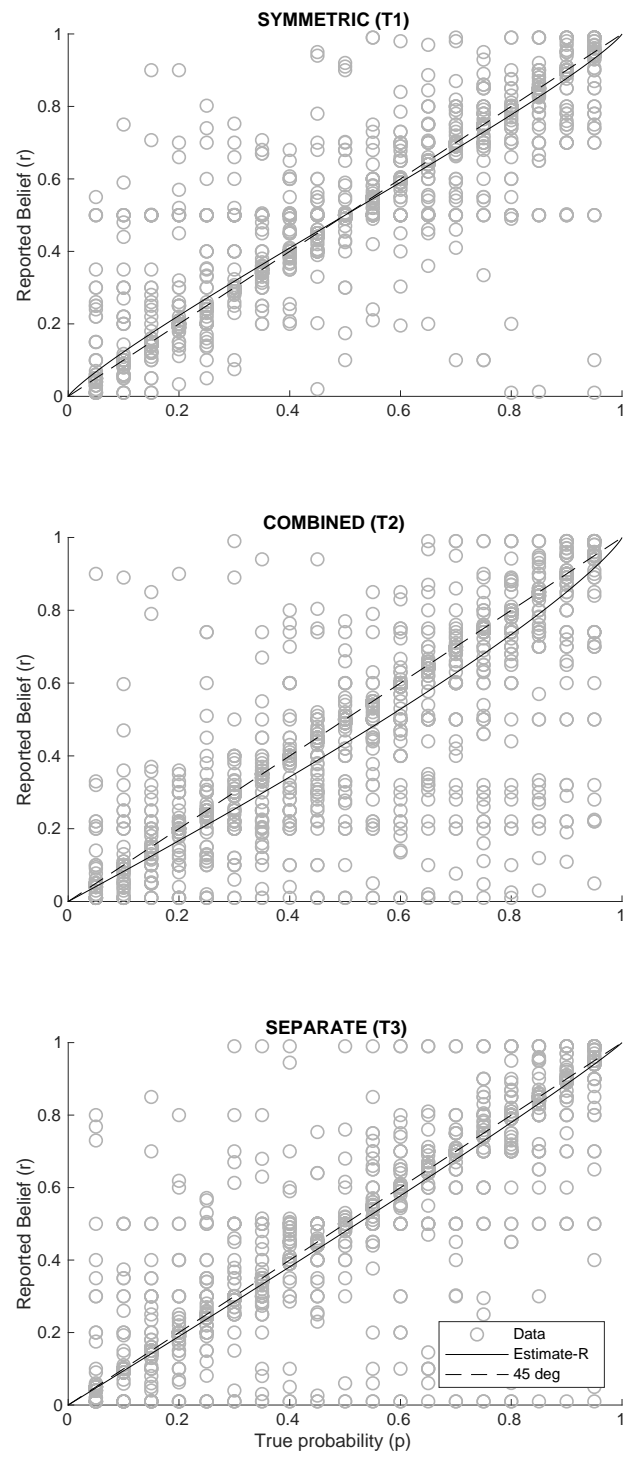
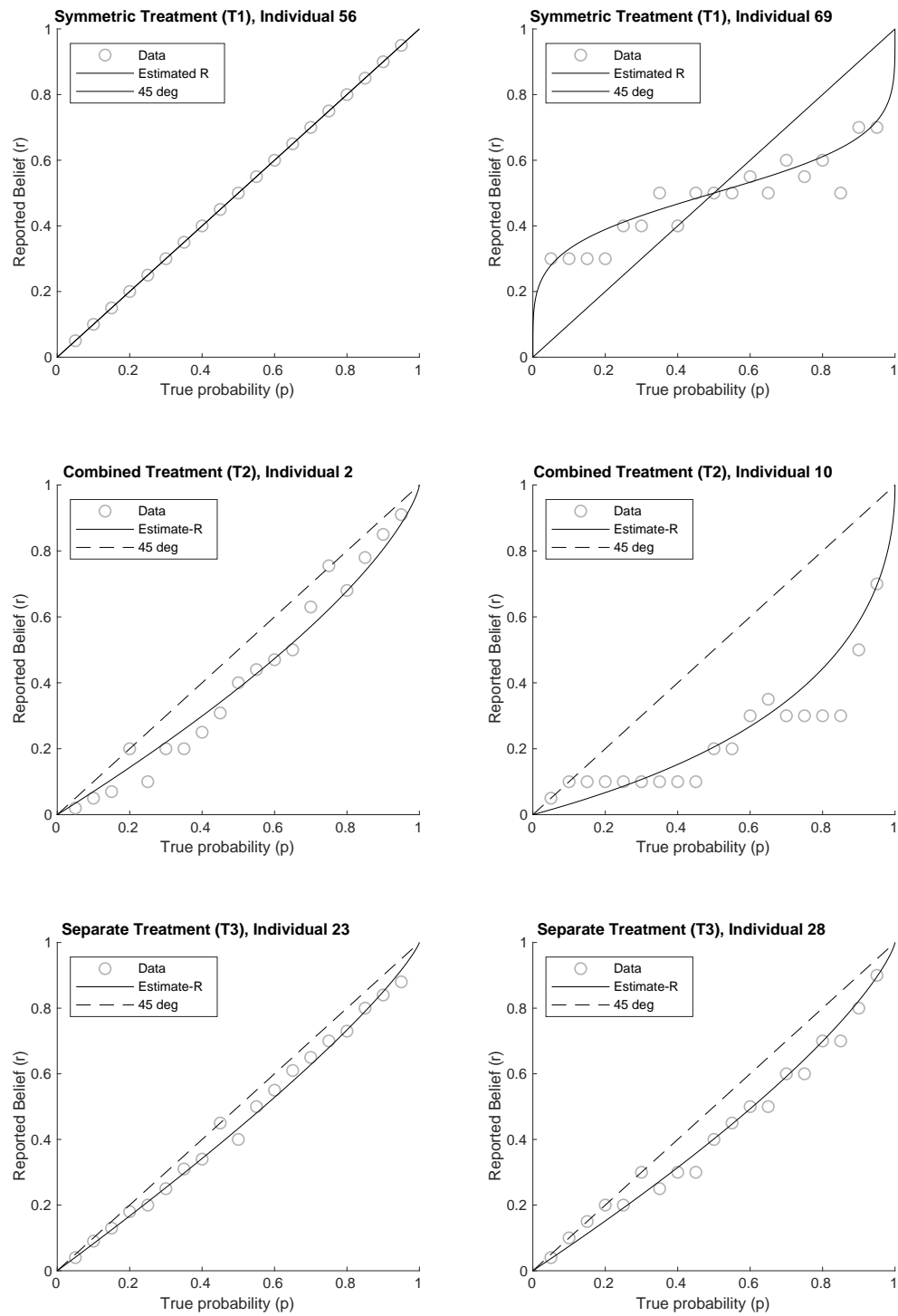


Figure 6: Individual Level Estimates of the Incentive Correction Function.



Appendix C: Endogeneity: An Illustrative Example

This section provides a simple illustrative example of why it is important to exercise caution in addressing possible endogeneity issues when studying belief updating. In particular, in many instances we are interested in studying how individuals update their beliefs from home-grown prior beliefs (i.e. subjective prior beliefs that are not exogenously endowed to subjects). However, when studying how subjects update beliefs from a home-grown prior, it is important to pay careful attention to the possible endogenous relationships between: (i) individual updating types, (ii) states of the world, and (iii) prior beliefs.⁶² In the illustrative example I discuss below, I consider the implications of a relationship between (i) and (ii), allowing for (iii) to be completely exogenous.

One domain where studying belief updating from home-grown priors is essential is the domain of beliefs about the *self*. Beliefs about one’s self deserve special attention since they are of critical importance in guiding our interaction with the world around us. This set of beliefs are also of central importance to the literature considering the *good-news*, *bad-news* hypothesis, since these beliefs are often heavy in affect, and are amongst the beliefs that we care most about.

However, when we study belief updating from subjective priors, there is a danger that an individual’s prior beliefs are related to the way she updates her beliefs. Furthermore, if the beliefs pertain to her *self*, her prior belief may be related to this fundamental, implying that the distribution of signals she receives is related to this fundamental and to her prior. This is true even if the individual receives “exogenous” noisy signals about the fundamental.

The following discussion has the objective of illustrating one possible way in which neglecting to pay attention to the endogeneity of the signal distribution can be problematic. The example is purely hypothetical, and rather contrived, but serves to illustrate the basic point. In particular, I use a very simple “toy” simulation to demonstrate that ignoring this issue can (in principal) lead to mistakenly find evidence for asymmetric updating when all individuals update symmetrically. It is important to point out that I am *not* suggesting that this is the explanation for asymmetric updating results observed in the literature⁶³ - I am simply highlighting a potential issue that should be addressed in this literature going forwards. In this regard, I also suggest a simple solution for dealing with the issue.

Much of the belief updating literature considers situations that resemble the following basic structure: consider an agent who updates about two states of the world, $\omega \in \{High, Low\}$, and receives a sequence of noisy signals, $s_t \in \{UP, DOWN\}$. This also reflects the setup considered in the

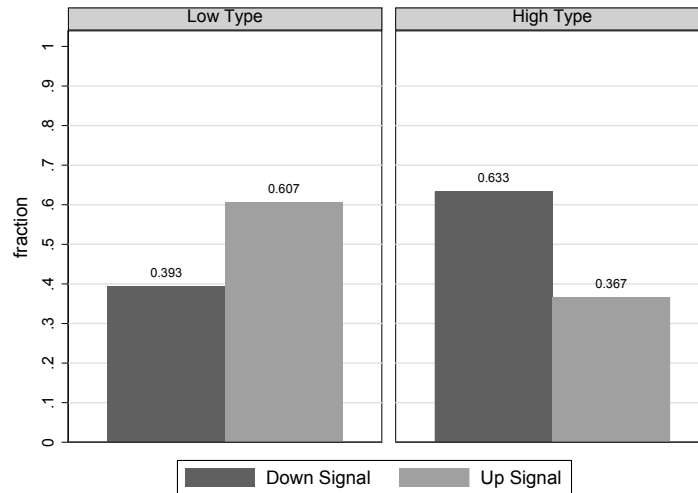
⁶²In some cases, e.g. when forming beliefs about the self, there are differences across individuals in the fundamental (ii) that they are forming beliefs about. In these cases, the relationship between (i) and (iii) might be mediated by a natural relationship between the fundamental (ii) and the priors.

⁶³For example, Figure 3 in Möbius et al. (2014) suggests that this is probably not a major concern for their main results. However, it is still important to control for this potential endogeneity issue as a robustness check.

current paper. However, since it is important to also understand how we form beliefs about *self*, in some studies the states are determined by personal characteristics of the individual (e.g. IQ). This means that states are essentially equivalent to personal types (i.e. *states* = *types*). The implication of this is that if signals are informative about the state of the world, then *High* types are more likely than *Low* types to receive *Up* signals (and vice versa for *Down* signals). If *High* types update their beliefs differently from *Low* types, this can (in principal) lead to finding (what looks like) evidence that the average individual updates asymmetrically when no individual actually does.

In order to show this, I conduct a very simple simulation exercise. I construct a population of 10 000 individuals who are randomly assigned to one of two types, $\omega \in \{High, Low\}$. Within each type, the agents' prior beliefs about the likelihood of being the *High* type are assigned randomly using a uniform distribution, distributed between zero and one.⁶⁴ *High* types receive an *Up* signal with probability $q = \frac{5}{8}$ and *Low* types receive a *Down* signal with probability $q = \frac{5}{8}$. Using a seed of 1000 in STATA, the observed empirical distribution of signals across types is shown in Figure 7.

Figure 7: Frequencies of Signals by Type



Now, the important part of this story is that belief updating may (in principal) be related to the underlying fundamental of interest. For example, it is conceivable that high IQ individuals process information and update their beliefs differently from low IQ individuals.⁶⁵

⁶⁴Note, this is an unrealistic assumption. In general, prior beliefs are related to the true state of the world. For example, beliefs about one's rank in an IQ distribution tend to be correlated with one's actual rank. However, for the purposes of this illustration, constructing type and prior belief to be orthogonal allows us to isolate only the effect of the endogeneity of types and signals (with exogenous priors). Allowing for priors to be related to the underlying fundamental would add an additional layer of endogeneity issues.

⁶⁵Note, even if one doesn't find the story that the two *types* (here, *High* and *Low*) might update their beliefs differently compelling, a very similar pattern could also be generated if there is a relationship between *prior beliefs* and

Here, we consider two types that use different, *but always symmetric*, updating rules. In particular, we consider a *High* type that is perfectly Bayesian, and a *Low* type that is not very responsive to new information (but otherwise very well behaved her belief updating).⁶⁶

The *High* type updates according to the following rule ($\delta = 1$, $\gamma_{UP} = 1$, $\gamma_{DOWN} = 1$):

$$\text{logit}(\pi_{t+1}) = 1 \cdot \text{logit}(\pi_t) + 1 \cdot \log\left(\frac{5}{3}\right) \cdot 1(s_{t+1} = UP) - 1 \cdot \log\left(\frac{5}{3}\right) \cdot 1(s_{t+1} = DOWN) \quad (20)$$

The *Low* type updates according to the following rule ($\delta = 1$, $\gamma_{UP} = 0.2$, $\gamma_{DOWN} = 0.2$):

$$\text{logit}(\pi_{t+1}) = 1 \cdot \text{logit}(\pi_t) + (0.2) \cdot \log\left(\frac{5}{3}\right) \cdot 1(s_{t+1} = UP) - (0.2) \cdot \log\left(\frac{5}{3}\right) \cdot 1(s_{t+1} = DOWN) \quad (21)$$

However, if we as the analyst neglect the possibility that the two types update their beliefs differently, then we might obtain biased parameters. This is illustrated by the regression estimates presented in column 1 and 2 of Table 7 below. These columns reflect the estimates from the standard specification used in this literature (i.e. equation 2). These parameter estimates, along with the true population averages are summarised as follows:

	True Parameter Values (Population Ave.)	Estimates
δ	1	1
γ_{UP}	0.6	0.7
γ_{DOWN}	0.6	0.5

It is clear from this that in spite of the fact that there is not a single individual in this population who updates asymmetrically that the estimated parameters suggest that there is an asymmetry. Notice, the standard errors are small and the adjusted R^2 suggests a good model fit (see Table

updating. This follows because: (i) prior beliefs about one's self are typically related to the underlying fundamental in question (e.g. beliefs about one's IQ positively correlated with actual IQ), (ii) types are mechanically related to the distribution of signals in the class of experiments we're considering, and therefore (iii) prior beliefs are related to the distribution of signals observed. Therefore, the story described in this section is worth paying attention to in any situation where at least one of the following might be violated:

- (1) states of the world \perp belief updating and (2) priors \perp belief updating

⁶⁶Note, neither of the types makes any errors in their belief updating. They both follow their updating rule perfectly. This exercise therefore rules out several other channels that can make life challenging for the analyst (e.g. errors related to priors or types).

7). As mentioned above, this is simply an illustration of why it is important to pay attention to the relationship between the distribution of signals and the types. Furthermore, it is important to point out that one can make a similar argument to show that even if the majority of individuals in the population are asymmetric updaters that the neglect of a relationship between signals and types could (in principal) generate estimates that suggest symmetric updating.

Fortunately, this particular endogeneity issue is easy to deal with by simply considering updating behavior within each type (e.g. interacting the RHS variables of equation 2 with the Type dummy variable). This is illustrated in column 3 of Table 7. Notice, also, that simply including the Type dummy variable in the regression does not solve the problem (see column 2).

Table 7: Estimates of Simulated Data Parameters

	Model 1 (1)	Model 2 (2)	Model 3 (3)
δ	1.002 (0.001)	1.002 (0.001)	1.000 (.)
γ_{UP}	0.696 (0.005)	0.683 (0.007)	0.200 (.)
γ_{DOWN}	0.504 (0.006)	0.513 (0.006)	0.200 (.)
High Type (=1)		0.0114 (0.004)	
High Type (=1) * δ			0.000 (.)
High Type (=1) * γ_{UP}			0.800 (.)
High Type (=1) * γ_{DOWN}			0.800 (.)
N	10000	10000	10000
Adjusted R^2	0.99	0.99	1.00

(i) Standard errors in parentheses

(ii) Note: Std errors in column 3 missing due to perfect fit.

Appendix D: Supplementary Figures, Results and Experimental Instructions

Figure 8: Comparison of Initial (Period 0) Belief with the Exogenous Prior

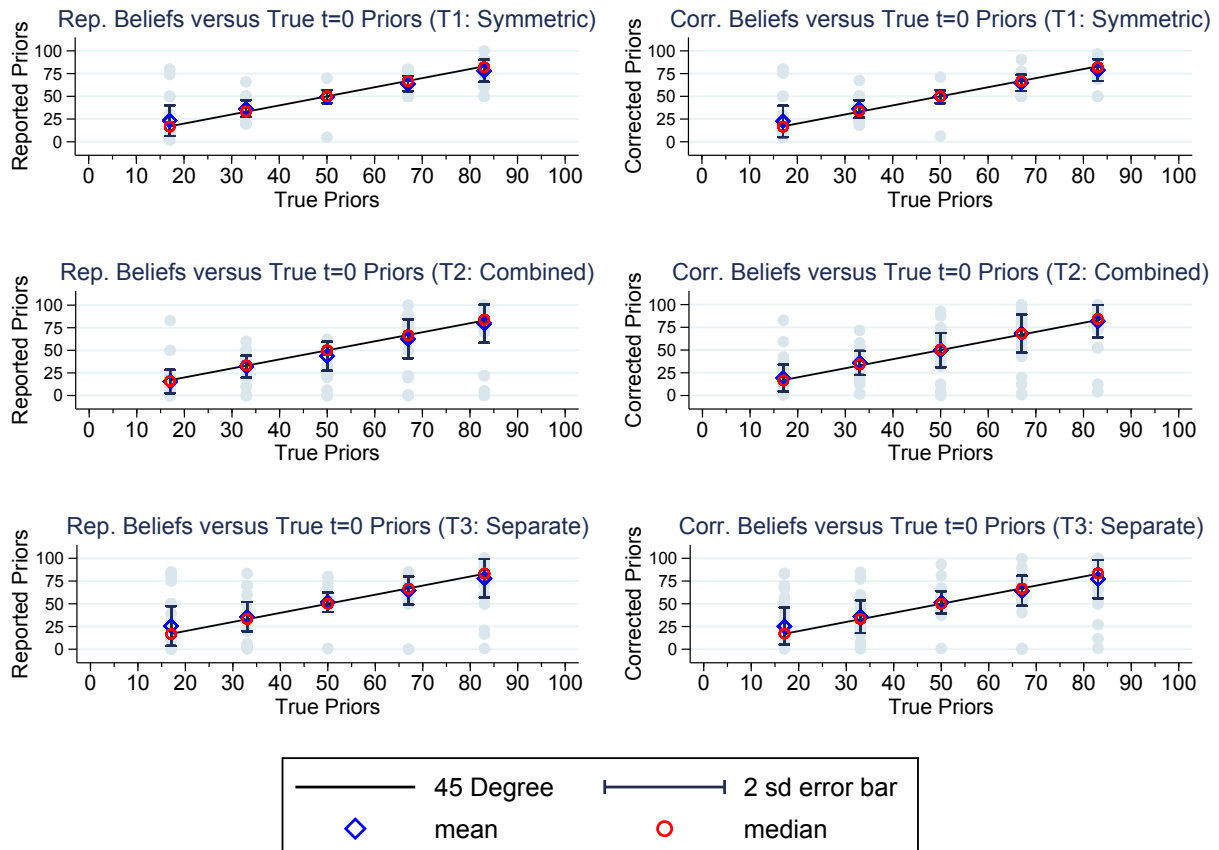


Table 8: First Stage Regressions Output Associated with Table 3

	<u>T1: SYMMETRIC</u>		<u>T2: COMBINED</u>		<u>T3: SEPARATE</u>	
	Reported (1a)	Corrected (1b)	Reported (2a)	Corrected (2b)	Reported (3a)	Corrected (3b)
δ_{IV}	0.81 (0.06)***	0.82 (0.05)***	0.95 (0.06)***	0.92 (0.06)***	0.88 (0.06)***	0.85 (0.05)***
γ_a	-0.07 (0.11)	-0.07 (0.10)	-0.60 (0.22)***	-0.05 (0.20)	-0.03 (0.25)	-0.09 (0.25)
γ_b	-0.17 (0.10)*	-0.20 (0.11)*	0.34 (0.20)*	-0.20 (0.17)	-0.24 (0.23)	-0.21 (0.26)
Kleibergen Paap F	170.39	228.03	238.80	256.15	220.64	276.53
Shea Partial R^2	0.58	0.61	0.48	0.50	0.50	0.49
N	1,075	1,075	1,285	1,285	1,140	1,140

(i) Standard errors in parentheses (clustered at the individual level).

Table 9: Testing for Sample Balance Across Treatment Groups

	T1 SYMMETRIC	T2 COMBINED	T3 SEPARATE	D(2,1)	D(3,1)	D(3,2)
Gender (Male = 1)	0.56 (0.50)	0.58 (0.50)	0.58 (0.50)	0.02	0.02	0
Age	22.88 (3.46)	23.02 (4.47)	21.88 (3.11)	0.14	-0.99	-1.13
Location (London = 1)	0.51 (0.50)	0.48 (0.50)	0.52 (0.50)	-0.03	0.01	0.04
Economics Class	0.44 (0.50)	0.54 (0.50)	0.42 (0.50)	0.10	-0.02	-0.12
Home Language German	0.42 (0.50)	0.45 (0.50)	0.42 (0.50)	0.03	-0.00	-0.03
Home Language English	0.28 (0.45)	0.20 (0.40)	0.38 (0.49)	-0.08	0.10	0.18**
Cognitive Reflection Score	2.07 (0.94)	2.02 (0.93)	2.05 (1.05)	-0.05	-0.02	0.03
<i>N</i>	57	65	60			

(i) Standard deviations in parentheses

(ii) T-test for difference being significant: * = 10%, ** = 5%, *** = 1%

Table 10: Average Updating Behavior across Treatments (Full Sample)

	<u>T1: SYMMETRIC</u>		<u>T2: COMBINED</u>		<u>T3: SEPARATE</u>	
	Reported (1a)	Corrected (1b)	Reported (2a)	Corrected (2b)	Reported (3a)	Corrected (3b)
<u>OLS</u>						
δ	0.74 (0.05)***	0.73 (0.06)***	0.80 (0.03)***	0.80 (0.03)***	0.84 (0.04)***	0.86 (0.03)***
γ_a	0.95 (0.10)	0.92 (0.09)	0.86 (0.10)	0.91 (0.10)	1.00 (0.11)	0.93 (0.10)
γ_b	0.83 (0.10)*	0.78 (0.10)**	0.99 (0.14)	0.90 (0.12)	0.96 (0.10)	0.91 (0.09)
$p(H_0 : \gamma_a = \gamma_b)$	0.26	0.19	0.25	0.90	0.64	0.80
N	1,875	1,875	1,850	1,850	1,825	1,825
R^2	0.52	0.51	0.62	0.64	0.67	0.70
<u>IV</u>						
δ	0.99 (0.03)	1.00 (0.03)	0.96 (0.02)	0.95 (0.02)**	0.99 (0.02)	1.00 (0.02)
γ_a	0.87 (0.09)	0.84 (0.09)*	0.87 (0.10)	0.83 (0.09)*	0.95 (0.11)	0.88 (0.09)
γ_b	0.82 (0.09)*	0.78 (0.09)**	0.89 (0.12)	0.89 (0.12)	0.92 (0.09)	0.88 (0.08)
$p(H_0 : \gamma_a = \gamma_b)$	0.47	0.44	0.83	0.51	0.74	0.98
N	1,875	1,875	1,850	1,850	1,825	1,825
1st Stage F	70.89	82.63	82.83	82.45	54.72	59.90

(i) Standard errors in parentheses (clustered at the individual level).

(ii) All coefficients are significantly different from 0 at the 1% level. Therefore, t-tests of the null hypothesis (H_0 : Coefficient = 1) are reported: * = 10%, ** = 5%, *** = 1%.

(iii) The rows corresponding to $p(H_0 : \gamma_a = \gamma_b)$ report the p-statistic from a t-test of the equality of the coefficients γ_a and γ_b (i.e. a test of the asymmetric updating hypothesis).