Learning from unknown information sources*

Yucheng Liang †

March 19, 2020

Abstract

When an agent receives information generated by a source whose accuracy might either be high or low, standard economic theory dictates that she update as if the source has medium accuracy. In a lab experiment, I find that subjects' updating behaviors deviate from this benchmark. First, subjects under-react to information when the source is uncertain. Second, the under-reaction is more pronounced for good news than for bad news. These two patterns, *under-reaction* and *pessimism*, are consistent with a theory of belief updating where agents are insensitive and averse to compound uncertainty and ambiguity. I also find that subjects' reactions to information with uncertain accuracy are uncorrelated with their evaluations of bets with uncertain odds. This suggests that people have distinct attitudes toward uncertainty in information accuracy and uncertainty in economic fundamentals. The experimental results are validated using observational data on stock price reactions to analyst earnings forecasts, where analysts with no forecast records are classified as uncertain information sources.

***Keywords*—** Belief updating, ambiguity, compound risk, earnings forecasts

*This study is approved by the Stanford IRB in Protocol 45930 and by the UCSB IRB in Protocol 2-18-0331. All errors are mine. Acknowledgements to be added. First version: Jan 13, 2019.

†Graduate School of Business, Stanford University. Email: ycliang@stanford.edu.

1 Introduction

People often need to incorporate new information for decision-making when they are uncertain about the accuracy of its source. For example, in response to a financial report issued by a new analyst, investors need to decide how to adjust their portfolios, but they don't know the analyst's expertise. When a patient hears about a health tip through the grapevine, she needs to decide whether to follow the tip, but it could either originate from a doctor or simply be an unsubstantiated rumor. Politicians often have to rely on media and polling agencies to learn about their constituents' needs, but they may not know the exact biases of these intermediaries.

Standard economic theory assumes that when the accuracy of information is uncertain, agents are able to correctly deduce *expected accuracy* and update their beliefs solely based on that expectation. As an example, consider a bet with two possible outcomes. The agent receives a prediction of the outcome but she does not know its accuracy: the probability that the prediction is correct conditional on the true outcome might either be 90% or 50%. The two possible levels of accuracy are equally likely, and the true one is independent from the realization of the outcome. According to standard theory, the agent is able to calculate that the prediction is correct 70% of the time in expectation, and her belief about the bet should react to this prediction as if she knew with certainty that it is 70% accurate.

1.1 Main results

Using both experimental data from the lab and observational data on stock price reactions to analyst earnings forecasts, this paper provides the first evidence on how uncertainty in information accuracy affects belief updating. In part of a lab experiment, I present subjects with bets and inform them about the winning odds. I elicit subjects' certainty equivalents (CEs) for each bet after they receive a prediction of its outcome. I refer to a prediction as *uncertain information* if its accuracy might either be high (ψ_h) or low (ψ_l), and I define its corresponding *simple information* as a prediction whose accuracy is known to be the midpoint, $(\psi_h + \psi_l)/2$. When subjects receive uncertain information, sometimes they know that the two possible accuracy levels are equally likely (*compound information*), and other times they don't know their relative likelihood (*ambiguous information*). In my experiment, the effects of compound uncertainty and ambiguity turn out to be qualitatively similar, so I will refer to them jointly using the generic term "uncertainty" and will not distinguish between them until the end of this section.

The main experimental result concerns the marginal effects of uncertainty in information accuracy on posterior beliefs. Compared to the case of simple information, subjects' beliefs move *less* in the direction of the realized message when its accuracy is uncertain. In other words, uncertain information leads to *under-reaction*. Moreover, reactions are more biased toward the direction of bad news. This suggests that uncertainty in information accuracy, on average, leads to *pessimism* in posterior beliefs.

The same patterns also emerge when I examine stock market reactions to analyst earnings forecasts. The forecast accuracy of financial analysts is more unpredictable if they don't have a proven forecast record for the stock. I find that in response to good news (upward forecast revisions) issued by these analysts, the immediate stock price reactions are, on average, followed by larger positive price drifts. This phenomenon implies that investors' immediate reactions to good news are less sufficient when the news is issued by analysts without records. In contrast, the sufficiency of reactions to bad news does not depend on whether the issuing analyst has a forecast record or not. These findings are consistent with my experimental results that people's reactions to information from unknown sources are less sufficient and more pessimistic, and they show that these phenomena are present even in a high-stake, real-world environment.

1.2 Nature of the behavioral patterns

A related question which has been extensively studied is how people evaluate prospects when they don't know the probability over payoff-relevant states.[1] For example, investors may need to evaluate a complex financial asset when the distribution of its returns is difficult to know. The answer to this question can shed light on the nature of people's reactions to information from unknown sources because in both problems, decision-makers face uncertainty about the probability over the state space. Following a canonical experimental design in this literature,[2] in a separate part of my experiment, I elicit subjects' CEs of *uncertain bets*, which are bets whose winning odds might either be high (p_h) or low (p_l). The results exhibit two patterns. First, compared to simple bets whose odds are known to be the midpoint, $(p_h + p_l)/2$, evaluations of uncertain bets change less as p_h and p_l change. The literature refers to this pattern as *uncertainty-induced insensitivity*, and it captures the psychological intuition that people internalize the odds less as they become more complex. Second, the lower winning probability is weighted more in the evaluation of uncertain bets than the higher one. This pattern is called *uncertainty aversion*, and it reflects the tendency to be over-pessimistic in face of compound uncertainty or ambiguity. These two patterns are empirical regularities in people's uncertainty attitudes and have been subjects of theorization.[3] To

[1] The question was raised in Keynes (1921), Knight (1921), and Ellsberg (1961), and has since received immense theoretical attention. For theoretical surveys, see Machina and Siniscalchi (2014) and Gilboa and Marinacci (2016). Trautmann and van de Kuilen (2015) provides a survey of empirical evidence.

[2] See e.g. Halevy (2007); Chew et al. (2017).

[3] Empirical studies that show evidence for uncertainty-induced insensitivity and uncertainty aversion include Abdellaoui et al. (2011, 2015); Dimmock et al. (2015); Baillon et al. (2018); Anantanasuwong et al. (2019). Theoretical models that capture these two patterns include Ellsberg (2015); Chateauneuf et al. (2007); Gul and Pesendorfer (2014).

parsimoniously captures uncertainty-induced insensitivity and uncertainty aversion, I introduce and axiomatize an *ε-α-maxmin expected utility* model of preference. This model is also convenient for structural estimation, which helps organize the experimental results.

The problem of learning from unknown information sources differs from evaluating bets with uncertain odds in two ways. First, it involves the arrival of new information. Second, the uncertainty is in the accuracy of information sources as opposed to economic fundamentals. Both differences could potentially alter the effects of uncertainty on behavior.

Uncertainty attitudes in problems with belief updating With the arrival of new information, uncertainty attitudes could manifest themselves in a number of different ways. For example, consider the main part of my experiment where subjects evaluate a bet after receiving a prediction whose accuracy might either be high or low. Suppose a subject is averse to uncertainty in information accuracy. Then, one possibility is that the uncertainty aversion leads her to be pessimistic about the outcome of the bet conditional on the prediction. (The belief updating rule that leads to this possibility is known as *Full Bayesian updating*,[4] which is different from Bayesian updating in the classical sense.) Alternatively, the subject may be pessimistic about the ex-ante value of information (*Dynamically consistent updating*).[5] A third possibility is that after hearing the prediction, the subject no longer considers one of the accuracy levels because she deems it unlikely (*Maximum likelihood updating*).[6]

I derive the empirical implications of these (and other) possibilities, and Full Bayesian updating coupled with uncertainty-induced insensitivity and uncertainty aversion stands out as being most in line with the experimental and stock market evidence I describe above. Intuitively, under Full Bayesian updating, uncertainty averse agents over-weight the possibility that uncertain information has low accuracy when they hear good news. On the contrary, the high accuracy is over-weighted when the realized news is bad. This asymmetry leads to pessimism about the value of the bet after hearing the news. On the other hand, uncertainty-induced insensitivity causes agents to neglect the content of information from unknown sources, leading to under-reaction. In addition, Full Bayesian updating also has predictions on belief updating with uncertain priors and simple information, and the results from another part of my experiment are broadly consistent with those predictions.[7]

[4] Jaffray (1992); Pires (2002); Eichberger et al. (2007).

[5] Hanany and Klibanoff (2007).

[6] Dempster (1967); Shafer (1976); Gilboa and Schmeidler (1993).

[7] Full Bayesian updating predicts that if an agent is insensitive and averse to uncertainty in priors, then when she updates from uncertain priors, she will under-weight them ("base-rate neglect") and be pessimistic about the bet's value. I find evidence that uncertain priors lead to more pessimism but not more base-rate neglect. This is perhaps because base-rate neglect is already severe even when priors are known.

Attitudes toward uncertain information accuracy vs. attitudes toward uncertain economic fundamentals People's attitudes toward uncertain information accuracy may be different from their attitudes toward uncertain economic fundamentals. To investigate this possibility, I compare lab subjects' reactions to uncertain information and their evaluations of bets with uncertain odds, both in the aggregate and at the individual level. In the aggregate, I estimate a representative-agent model based on the ε-α-maxmin preference and Full Bayesian updating. The estimates show that the representative agent is insensitive and averse to both uncertainty in information accuracy and uncertainty in the odds of bets. The degrees of insensitivity and aversion are also similar across different kinds of uncertainty.[8]

One might be inclined to conclude from the aggregate results that uncertainty attitudes toward information sources and economic fundamentals are similar. However, the similarity completely breaks down when we focus on individual subjects. At the individual level, I construct tests for the correlations between attitudes toward different kinds of uncertainty. These tests are valid under a variety of preference models and updating rules. The results show that there is almost zero correlation between attitudes toward uncertainty in information accuracy and uncertainty in the odds of bets. This stark finding suggests that knowing a person's preference between simple and complex assets does not help predict how she reacts differently to information from known and unknown sources. This result resonates with existing evidence showing that uncertainty attitudes vary with issues and are susceptible to framing (e.g., Heath and Tversky, 1991; Fox and Tversky, 1995).

1.3 Compound information vs. ambiguous information

So far, I have not distinguished between compound and ambiguous information. While both kinds of information are complex, only ambiguous information has the additional feature that the probability distribution over possible accuracy levels is not explicitly specified. Since in real life, most information sources exhibit varying degrees of both features, comparing the effects of compound and ambiguous information helps disentangle the roles of complexity and "unknown unknowns" in generating belief updating biases.

At the aggregate level, structural estimation using the experimental data shows that the representative agent is insensitive and averse to both compound uncertainty and ambiguity. Moreover, the magnitudes of insensitivity and aversion are larger for ambiguous information than for compound information. These results suggest that both complexity and "unknown unknowns" play a role in the effects of uncertain information on belief updating. At the individual level, in more than one third of all cases, a subject has the same reaction to compound and ambiguous information.[9] Taken

[8]One exception is that the degree of insensitivity to ambiguous odds of bets is imprecisely estimated in belief updating tasks with ambiguous priors and simple information.

[9]Similar results also hold for the relationship between compound and ambiguous economic fundamentals,

together, these findings imply that in order to mitigate the belief updating biases induced by uncertain information accuracy, the meanings of messages should not only be precisely specified, but also be presented in a simple way.

1.4 Related literature

Theoretical studies have proposed various criteria for belief updating under compound uncertainty and ambiguity (Dempster, 1967; Shafer, 1976; Jaffray, 1992; Epstein and Schneider, 2003; Hanany and Klibanoff, 2007; Cheng and Hsiaw, 2018; Gul and Pesendorfer, 2018). Many theories differ in their predictions on the marginal effects of uncertainty on posterior beliefs, which allows me to empirically distinguish between them. Several experimental papers have studied learning with uncertain priors. Corgnet et al. (2012), Ert and Trautmann (2014), and Moreno and Rosokha (2016) study choices between bets after sampling from urns with uncertain compositions.[10] They find that beliefs converge to the true distribution with sampling, but results on the rate of convergence and the evolution of ambiguity attitudes are mixed. Ngangoué (2018) elicits CEs of ambiguous bets with and without additional simple information in a between-subject design, and she finds support for recursive smooth preferences (Klibanoff et al., 2009). In contrast to these earlier studies, the main focus of my paper is on uncertain information accuracy.[11]

Two previous experimental projects study phenomena related to uncertain information accuracy. Fryer et al. (2018) find that subjects tend to update their beliefs about political issues in the directions of their priors after reading ambiguous research summaries. In a social learning experiment, De Filippis et al. (2018) present subjects with two pieces of information: a private signal about the true state and the belief of a predecessor (who only has a private signal). When the private signal is absent or confirms the predecessor's belief, subjects account for the predecessor's belief in a Bayesian manner. In contrast, when the private signal contradicts the predecessor's belief, subjects under-weight the latter. The authors interpret their result using a model where subjects treat their predecessors' beliefs as ambiguous information. My experiment differs from these two studies in that I examine the effects on belief updating when information accuracy changes from being *objectively* simple to *objectively* uncertain. In addition, the context of my experiment rules out explanations that resort to motivated reasoning.

which confirm previous findings in Halevy (2007); Abdellaoui et al. (2015); Chew et al. (2017); Gillen et al. (forthcoming).

[10]In these papers, the (simple) information pertains to the true composition of the urn, which is different from my experiment where the information is about the realized outcome ("the ball drawn from the urn"). For a similar design with natural events instead of artificial urns, see Baillon et al. (2017).

[11]Cohen et al. (2000) and Dominiak et al. (2012) study dynamic Ellsberg three-color experiments where subjects choose between bets before and after one color is ruled out. They find that most subjects' behavior is consistent with Full Bayesian updating, though their experiments do not elicit point beliefs.

Two recent experimental studies investigate certain aspects of ambiguous information. In a contemporaneous project, Epstein and Halevy (2019) study belief updating with ambiguous information when the prior is compound. Using a between-subject design, they find that more subjects violate the martingale property of belief updating[12] under ambiguous information than under a piece of simple information (which is not the symmetric reduction of the ambiguous information). More recently, Shishkin ⓡ Ortoleva (2019) focus on ambiguous neutral information (i.e. information whose accuracy is a midpoint-preserving spread of 50%) and study both belief updating and the demand for information. In contrast to these two studies, my experiment allows for the separate identification of under-reaction and pessimism caused by uncertain information accuracy. Also, I consider both compound and ambiguous information.

Two empirical studies find patterns in real-world settings which can be explained by certain models of learning from ambiguous information. Epstein and Schneider (2008) calibrate the US stock price movement in the month after 9/11 to a model of asset pricing with ambiguous news and find that the fit is superior to a Bayesian model. Kala (2017) studies how rainfall signals affect Indian farmers' agricultural decisions and find support for Hansen and Sargent (2001)'s model of robust learning. These papers do not study how the degree of uncertainty in information accuracy affects the sufficiency of reactions to news, which is what I focus on in the analysis of stock price reactions to analyst earnings forecasts.

There is a vast literature on stock market reactions to analyst reports in accounting and finance.[13] Gleason and Lee (2003) find that stock price under-reaction is less pronounced for analysts who are recognized by the *Institutional Investor* magazine. Zhang (2006) shows that the market under-reacts more to forecast revisions on firms whose fundamentals are harder to learn. Complementary to these studies, my paper focuses on the uncertainty of analysts' accuracy, and I find that it exacerbates under-reaction only for good news. Mikhail et al. (1997) and Chen et al. (2005) study how analysts' experience and forecast records affect the market's immediate reactions to their forecasts, but they don't study the sufficiency of these reactions.

More broadly, my paper is related to the fast-growing literature on belief updating biases, such as under-reaction (e.g., Edwards, 1968; Möbius et al., 2014) and asymmetric updating (e.g., Eil and Rao, 2011; Möbius et al., 2014; Coutts, 2019; Barron, 2019). Benjamin (2019) surveys this literature and concludes that evidence on the directions of belief updating biases is mixed. Although most experimental studies on these topics focus on people's reactions to objectively simple information, due to inattention or bounded rationality, people may still perceive the information as uncertain to varying degrees. If this is true, then my paper suggests that perceived uncertainty in information accuracy may be a moderator for these belief updating biases. Indeed, in a subsequent experiment,

[12]Loosely speaking, the martingale property of belief updating states that there exists a probability distribution over messages such that for every event, the expectation of posteriors equals the prior.

[13]For surveys, see Kothari et al. (2016); Bradshaw et al. (2017).

Enke and Graeber (2019) find evidence that supports this conjecture.[14]

1.5 Paper structure

The rest of the paper is organized as follows. Section 2 describes the design of all parts of the lab experiment. Section 3 presents experimental results on the evaluation of bets without new information and introduces the ε-α-maxmin EU model of uncertainty attitudes. The main experimental results are in Section 4. First, I discuss theories of learning with uncertain information accuracy that combine the ε-α-maxmin model with a belief updating rule. Then, to test these theories, I present the experimental results on belief updating with uncertain information. As a validation, in Section 5, I briefly go through the implications of these belief updating rules in a setting of belief updating with uncertain priors and summarize the corresponding experimental evidence. To measure and compare the aggregate-level effects of different kinds of uncertainty, I introduce and estimate the structural model in Section 6. The individual-level relation between attitudes toward uncertain information accuracy and uncertain economic fundamentals is analyzed in Section 7, whereas the relation between attitudes toward compound uncertainty and ambiguity is discussed in Section 8. Section 9 presents supporting evidence using observational data on stock market reactions to analyst earnings forecasts. Section 10 concludes.

2 Experimental design

2.1 Tasks

Each session of the experiment consists of 29 rounds. Each round is framed as a race between a Red horse and a Blue horse with no tie. In each round, there are two payoff-relevant states, *Red* and *Blue*, corresponding to the color of the winning horse. A *bet* on a state pays out \$20 if it is the true state, and \$0 otherwise. At the end of each round, subjects report their certainty equivalents (CEs) of the Red bet and the Blue bet in an incentive-compatible mechanism. Before reporting their CEs, subjects may receive a piece of additional information, framed as an analyst report, about the true state. The message they receive is either "Red horse won" (*red*) or "Blue horse won" (*blue*). The uncertainty across rounds is independent.

The 29 rounds are grouped into 5 parts, which are summarized in Table 2.1. In the three parts with "simple priors", subjects know with certainty the prior probability distribution over the states, i.e. the winning odds of the two horses. For example, subjects may be told that the Red horse has

[14] Enke and Graeber (2019) find in a part of their experiment that subjects' posterior beliefs are more compressed toward the prior (50%, 50%) if the information accuracy is compound as opposed to simple. Under their belief elicitation mechanism, this phenomenon is consistent with both under-reaction and pessimism.

Part	Prior	Information accuracy
1	Simple	No information
2	Simple	Simple
3	Simple	Uncertain
4	Uncertain	No information
5	Uncertain	Simple

Table 2.1: Summary of parts

70% chance of winning and the Blue horse has 30% chance. What differ across these three parts are whether subjects receive additional information about the true state before they report their CEs for the bets and, if they do, whether the accuracy of information source is uncertain. In Part 1, subjects do not receive additional information, whereas in Parts 2 and 3, they do. In Part 2, subjects know the accuracy level of information, denoted by ψ, with certainty. For instance, subjects may be told that the analyst report they receive is 70% accurate. This means that conditional on the true outcome of the horse race, the analyst report is correct 70% of the times and wrong 30% of the times. In Part 3, subjects know that the information is at one of two possible accuracy levels, ψ_h or ψ_l ($\psi_h > \psi_l$), but don't know which one. For example, they may be told that the analyst report is either 90% accurate or 50% accurate. In half of the rounds, subjects know that the two possible accuracy levels are equally likely to be the true one. I refer to this situation as subjects receiving *compound information*. In the other rounds, the distribution over the two possible accuracy levels is unknown, and I will refer to the information as *ambiguous information*. The realization of the true accuracy level is independent from the realization of the state.

In Parts 4 and 5, subjects are told in each round that the states are distributed according to one of two possible priors. For example, the prior probability that *Red* is the true state is either 50% or 90%. In half of the rounds, subjects know that the two possible priors are equally likely to be true ("compound prior"), whereas in the others, they don't know the distribution over them ("ambiguous prior"). Subjects don't receive additional information about the true state in Part 4, whereas in Part 5 they do. Moreover, the additional information in Part 5 is simple, i.e. its accuracy level is known with certainty.

Table 2.2 shows the priors and information that subjects may face in the experiment. For example, if in a round the prior of *Red* is 70% and the prior of *Blue* is 30%, then the prior of the round is written as (70%, 30%). The priors of the three rounds in Part 1 are (50%, 50%), (60%, 40%) and (70%, 30%). In Part 4, subjects are presented with the corresponding compound and ambiguous priors listed in the second column. In each of the other parts, subjects face four possible combinations of priors and information accuracy, both for compound uncertainty and for ambiguity. The four possible combinations correspond to the four rows in the table.

Simple prior	Uncertain prior	Simple information accuracy	Uncertain information accuracy
(50%, 50%)	(90%, 10%) or (10%, 90%)	70%	90% or 50%
(60%, 40%)	(90%, 10%) or (30%, 70%)	60%	90% or 30%
(70%, 30%)	(90%, 10%) or (50%, 50%)	50%	90% or 10%
		70%	90% or 50%

Table 2.2: Priors and information accuracy

2.2 Logistics

The experiment was conducted at the Econ Lab at University of California, Santa Barbara on May 9th and 14th-16th, 2018. A total of 165 subjects were recruited using ORSEE (Greiner, 2015) to participate in 11 sessions which lasted on average 90 minutes. Subjects watch instructional videos at the outset of the experiment and also before each part. After each video, screenshots and scripts are distributed to subjects in paper for their reference. Before proceeding to the first round of each part, subjects answer several comprehension questions to demonstrate that they understand the instructions. Both the videos and the comprehension questions take extra care to making sure that subjects understand the statistical meaning of priors and information accuracy, though there is no mention of any updating rules. The experiment ends with an unincentivized survey. The instructional videos, their scripts and sample screenshots of the rounds can be found on my website.

2.3 Payment

Each subject receives a \$5 show-up fee and, if they finish the experiment, a \$10 completion fee. The amounts of bonus they receive depend on their decisions in the experiment. Throughout the experiment, each subject reports CEs for $29 \times 2 = 58$ bets. To eliminate income effect, only one randomly selected bet counts for bonus. The CE for a bet is elicited using the Becker-DeGroot-Marschak (BDM) mechanism (Becker et al., 1964). Specifically, a price between \$0 and \$20 is randomly selected. If a subject's CE for the bet that counts for bonus is higher than the price, then her bonus will equal the payout of that bet. Otherwise, her bonus will equal the price. In the first two sessions, the original version of BDM mechanism was implemented and subjects were asked to write down their minimum selling prices for each bet on paper.[15] In the other 9 sessions, the BDM mechanism was implemented through a multiple price list programmed using oTree (Chen et al., 2016), where a subject makes a series of binary choices between receiving the bet and receiving a sure amount of money incrementing from \$1 to \$19 with a step of \$1. The CE is inferred to be the

[15] A total of 38 observations from 3 subjects in these two sessions are missing.

minimum sure amount that the subject chooses over the bet.[16] Throughout the 29 rounds, subjects do not receive any feedback.

2.4 Implementation of randomization

To encourage subjects to consider each bet and each price in isolation (Baillon et al., 2015) and also to establish the credibility of the random incentive mechanism, the randomization is conducted publicly before the first round of each session. Specifically, each subject draws two envelopes from two bags, one from each. One envelope contains the bet that will count for bonus and the other contains the price of the bet (row in the multiple price list).

In each round, each binary event is determined by a random draw from a deck of ten cards numbered from 1 to 10, one card for each number. To determine the true state, a small number on the drawn card corresponds to *Red* being the true state and a large number corresponds to *Blue*. The threshold number is determined by the true prior over the states. For example, suppose the true prior of *Red* is 70%. Then the Red horse wins if a number between 1 and 7 is drawn, and the Blue horse wins if the number is between 8 and 10. In rounds with additional information, the analyst report is correct if the number drawn from a second deck of cards is small, and wrong if the number is large. The threshold number corresponds to the true accuracy level of the report. Another deck of cards is used in rounds with two possible priors. If the two priors are equally likely, then which prior is the true prior depends on whether the draw from this deck is between 1 and 5. If the distribution over the two priors is unknown, then the threshold number that determines the true prior is not disclosed to the subjects.[17] When the information accuracy is uncertain, the true accuracy is determined in a similar fashion.[18] [19] After drawing the cards, the experimenter announces the realized message to the subjects. Then the subjects report their CEs for the Red bet and Blue bet. In the first two sessions, how the cards determine the events is explained in details to the subjects in the instructional videos. In the other sessions, only the determination of states is explained in details in the videos. Subjects are told that the realized message, true prior, and true accuracy level of the information are similarly determined by random draws from separate decks of cards. They are referred to the printed scripts for more details if they are interested.

[16] Multiple switching between the left and right sides of the list is not allowed.

[17] To mitigate the concern that the experimenter manipulates the threshold number ex post, subjects are told that the threshold number is printed on a paper and they are welcome to inspect it after the experiment.

[18] Instructions are framed such that the uncertainty about true prior or the uncertainty about the accuracy level of the information is always resolved first.

[19] In the first two sessions, to determine the true prior or the true accuracy level of information, a card is drawn from a deck of eight cards instead of ten. The uncertainty is resolved by whether the number drawn is even or odd.

2.5 Order between rounds

The order between the five parts varies across sessions. Within Parts 3, 4 and 5, the rounds with compound uncertainty are grouped in one block and the ones with ambiguity are grouped in another. The order between compound and ambiguous blocks also varies across sessions. Table B.2 summarizes the orders in the 11 sessions. In Appendix B.2, I show that order effects are not a driving force of the main empirical results.

Within each part (or each block in Parts 3, 4 and 5), the order between rounds is fixed, which is shown in Table B.3.

3 Evaluating bets with uncertain odds

Before discussing the main results of this paper on learning from unknown information sources, I first present experimental evidence on how people evaluate bets with unknown odds in a setting without belief updating. This evidence helps motivate a model of uncertainty attitudes, which is the basis of most theories of belief updating with uncertain information accuracy.

3.1 Model setup

Consider an agent choosing between a bet and a sure amount of utils. There are two payoff-relevant states, G and B. The bet pays out 1 util if G occurs and 0 util otherwise. I define the *evaluation* of the bet as the amount of utils u such that the agent is indifferent between the bet and u. Let p be the probability of G. Then, a standard expected utility (EU) agent will evaluate the bet by $u = p \cdot 1 + (1 - p) \cdot 0 = p$. If the state G has a compound probability, i.e. its probability is either p_h or p_l, each with equal chance, then a standard EU agent will evaluate the bet by $u = \frac{p_h + p_l}{2}$. The same evaluation also applies to the case of ambiguous probability if a standard EU agent treats p_h and p_l symmetrically under the principle of insufficient reason.

3.2 Experimental results

Figure 3.1 and Table 3.1 show the CEs of simple, compound and ambiguous bets in Parts 1 and 4 of my experiment where subjects do not receive additional information.[20] The mean CEs of simple

[20]Since the red and blue bets in a (50%, 50%) horse race are both bets with 50% chance of winning, I take the average of the CEs of the two bets to be the CE of a 50%-odds bet. In the simple round whose prior is (50%, 50%) and in its two corresponding uncertain rounds, 82% of the subjects report the same CE for the red and blue bets, which is in line with results in previous studies. See Table C. VI of Chew et al. (2017) for a meta-study. Moreover, the deviations from color neutrality are not significantly different from zero.

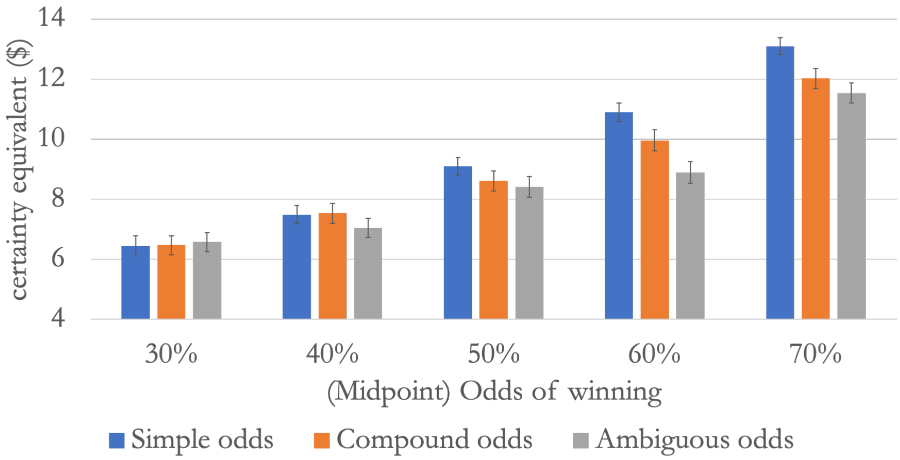

Figure 3.1: CEs of bets without additional information

Notes: The figure shows the mean CEs of bets without belief updating. Each group of bars represents the three tasks that share the same (midpoint) odds of winning. Error bars represent +/- one standard error.

bets are lower than their expected values except when the odds of winning is 30%. This is consistent with Prospect Theory (Kahneman and Tversky, 1979).

My main focus, however, is on the comparison between CEs of uncertain bets and simple bets. The mean CEs of uncertain bets are lower than their simple counterparts for medium and high odds. Nevertheless, the gaps vanish for bets with low odds (30% for ambiguous bets, 30% and 40% for compound bets). These results are consistent with two common patterns in people's attitudes toward compound uncertainty and ambiguity: *uncertainty aversion* and *uncertainty-induced insensitivity*.[21] With uncertainty aversion, the lower winning probability is weighted more in the evaluation of uncertain bets than the higher one. This reflects the tendency to be over-pessimistic in face of compound uncertainty or ambiguity. With uncertainty-induced insensitivity, evaluations of bets change less as the odds change if the odds are uncertain. This captures the intuition that people internalize the odds less as they become more complex.

3.3 A model of uncertainty attitudes

Previous studies have proposed a wide variety of models to account for uncertainty attitudes that deviate from standard expected utility theory.[22] In most parts of this paper, instead of trying to distinguish among these models, I will use a simple model that parsimoniously captures uncertainty

[21]For similar empirical patterns, see Abdellaoui et al. (2011, 2015); Dimmock et al. (2015); Baillon et al. (2018); Anantanasuwong et al. (2019).

[22]Gilboa and Marinacci (2016); Machina and Siniscalchi (2014) provide surveys of this literature.

(Midpoint) Odds of winning	Type of bet	Mean CE	Standard error	N
30%	simple	6.45	0.330	165
	compound	6.47	0.321	165
	ambiguous	6.57	0.313	165
40%	simple	7.49	0.300	165
	compound	7.53	0.331	165
	ambiguous	7.05	0.310	164
50%	simple	9.10	0.293	162
	compound	8.61	0.335	164
	ambiguous	8.41	0.343	163
60%	simple	10.89	0.308	165
	compound	9.96	0.355	165
	ambiguous	8.90	0.359	164
70%	simple	13.09	0.282	165
	compound	12.02	0.327	165
	ambiguous	11.53	0.333	165

Table 3.1: Bets without additional information

Notes: This table compares the mean CEs of bets with simple, compound and ambiguous odds without additional information. The last two columns are p-values for two-sided paired *t*-tests.

aversion and uncertainty-induced insensitivity.

There are two events, E and E^c. The probability of E is either p_h or p_l, with $p_h \geq p_l$. The probability of E^c is the complement. An act assigns a simple lottery to each event. I identify the simple lotteries assigned to E and E^c by their (von Neumann-Morgenstern) utility indices, denoted by u_1 and u_2, respectively. Define the following function:

$$W(x, y; \varepsilon, \alpha) := (1 - \varepsilon)[(1 - \alpha)x + \alpha y] + \varepsilon \cdot 0.5.$$

Then the utility of the act for an ε-α-maxmin agent is

$$\begin{cases} W(p_h, p_l; \varepsilon, \alpha) \cdot u_1 + (1 - W(p_h, p_l; \varepsilon, \alpha)) \cdot u_2, & \text{if } u_1 \geq u_2 \\ W(p_l, p_h; \varepsilon, \alpha) \cdot u_1 + (1 - W(p_l, p_h; \varepsilon, \alpha)) \cdot u_2, & \text{if } u_1 < u_2 \end{cases},$$

where ε and α are constants in $[0, 1]$. In the setting of my experiment where the payoff of the bet is either 1 util or 0 util, an agent with *ε-α-maxmin (expected utility) preference* evaluates the bet by

$$u = W(p_h, p_l; \varepsilon, \alpha)$$

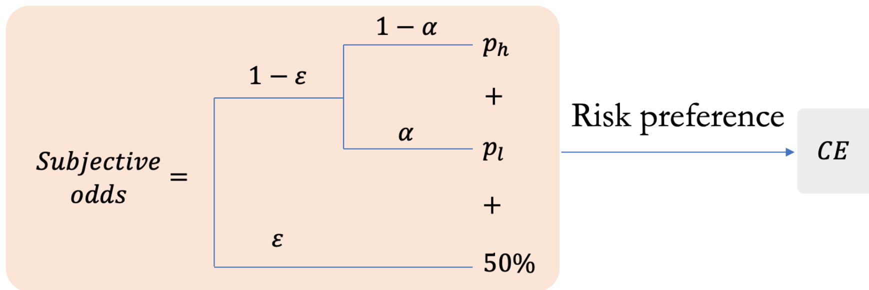

Figure 3.2: Illustration of the ε-α-maxmin EU preference

Notes: The figure illustrates how an ε-α-maxmin EU agent forms the CE of a bet with uncertain odds. Similar to a standard EU agent, she behaves as if she forms a subjective odds of the bet and then applies her risk preference to form the CE. Unlike a standard EU agent, the subjective odds of an ε-α-maxmin EU agent is a weighted average between p_h, p_l, and 50%.

and apply her risk preference to translate utils to certainty equivalents. (See illustration in Figure 3.2.)

The ε-α-maxmin preference can be interpreted as follows. When the probability distribution on the two events is uncertain, the agent puts ε weight on a baseline probability, which I assume to be the symmetric and maximally uncertain distribution (50%, 50%).[23] The rest of the weight is split between the more optimistic belief and the more pessimistic one in a generically asymmetric way. The two parameters flexibly capture rich patterns of attitudes toward uncertainty, from full uncertainty seeking ($\alpha = 0$) to full uncertainty aversion ($\alpha = 1$), from full sensitivity ($\varepsilon = 0$) to full insensitivity ($\varepsilon = 1$).

This model is closely related to many classes of models in the literature. It can be written in the functional form of Choquet expected utility (Schmeidler, 1989) (See Appendix C.1). Also, it generalizes the Maxmin expected utility preference (Gilboa and Schmeidler, 1989) and the α-maxmin expected utility preference (Olszewski, 2007). As a result, standard expected utility is also nested as a special case ($\varepsilon = 0$ and $\alpha = 0.5$). Similar two-parameter models have been proposed (Ellsberg, 2015; Chateauneuf et al., 2007) and fitted to data in experimental settings that are different from mine (Abdellaoui et al., 2011; Dimmock et al., 2015).[24] In Appendix D, I provide an axiomatic foundation of the ε-α-maxmin preference. I will also discuss two alternative models, the smooth

[23] I consider two alternative models in Appendix B.5. In one model, the weight on (50%, 50%) scales with $p_h - p_l$. In the other, the baseline probability on which the agent puts ε weight is a free parameter.

[24] In these model, agents face a continuum of possible probability distributions. Hence, the insensitivity parameter is often interpreted as capturing the range of distributions that agents deem possible. This interpretation does not fit my experimental setting as subjects are explicitly informed that an uncertain bet only has two possible winning probabilities.

model (Klibanoff et al., 2005) in Appendix E and Segal (1987, 1990)'s recursive model in Appendix F.

4 Belief updating with uncertain information accuracy

In this section, I analyze theories and experimental results on belief updating with uncertain information accuracy, which is the main subject of interest in this paper. The theories are based on the ε-α-maxmin EU model, which was introduced in the previous section.

4.1 Theories

As in the previous section, consider the choice between a sure amount of utils and a bet that pays out 1 util under state G. The probability of G is p with certainty. Before an agent makes the choice, she receives an additional piece of binary information $m \in \{g, b\}$. If the accuracy of the information is $Pr(g|G) = Pr(b|B) = \psi$ with certainty, then after observing the realized message, a Bayesian EU agent will evaluate the bet by the Bayesian posterior belief on G: $u(g) = Pr^{Bayes}(G|p, g, \psi) := \frac{p\psi}{p\psi+(1-p)(1-\psi)}$, $u(b) = Pr^{Bayes}(G|p, b, \psi) := \frac{p(1-\psi)}{p(1-\psi)+(1-p)\psi}$.

In a belief updating problem where the information accuracy is uncertain, the choice between the bet and a sure amount of utils conditional on information depends on the agent's uncertainty attitudes and her belief updating rule. In an *uncertain information* problem, the prior probability of G is still simple, but the accuracy of additional information is either ψ_h or ψ_l. The two levels of accuracy satisfy $0 < \psi_l < \psi_h < 1$ and $\psi_h + \psi_l \geq 1$. In the rest of this section, I will apply several belief updating rules to the ε-α-maxmin preference and compare their predictions. (See illustration in Figure 4.1.) Specifically, I will examine, under each updating rule, how choices given uncertain information deviate from those given simple information. I will also investigate how uncertainty attitudes (i.e. ε and α) affect these choices. Proofs of results in this subsection are in Appendix C.2.

4.1.1 Full Bayesian updating

In an uncertain information problem, Full Bayesian updating dictates that the evaluation of a bet conditional on good news is given by

$$u = Pr^{Bayes}(G|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha))$$

and that conditional on bad news is

$$u = Pr^{Bayes}(G|p, b, W(\psi_l, \psi_h; \varepsilon, \alpha)).$$

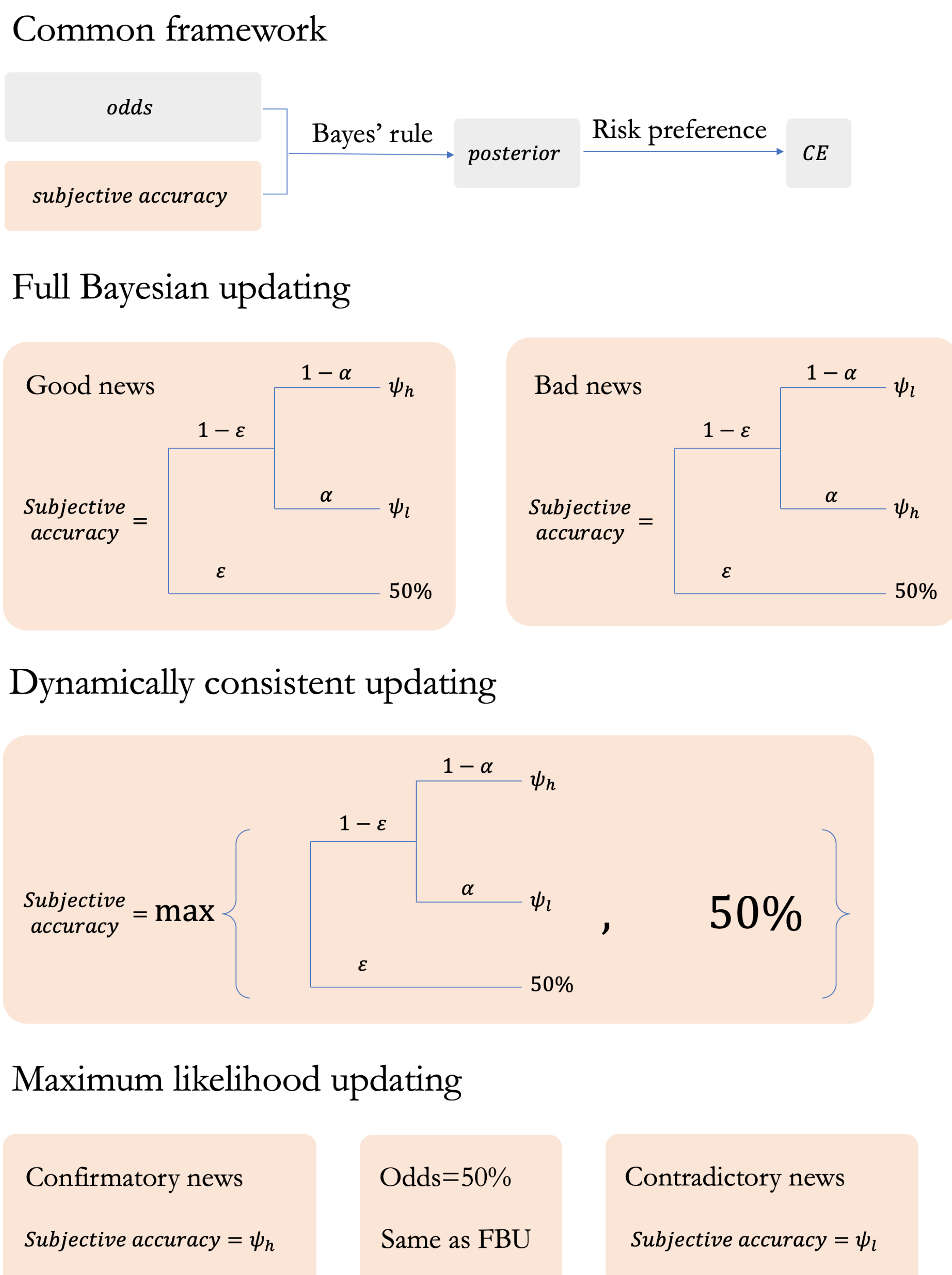

Figure 4.1: Illustration of theories of updating with uncertain information

These formulas can be derived from Eichberger et al. (2007) and they admit a simple interpretation. The agent behaves as if she perceives the information accuracy to be a weighted average of ψ_h, ψ_l and 50%, and she updates by applying the Bayes' rule to the prior and the subjective accuracy. The weight on 50% is always ε, which is responsible for the degree of under-reaction to news. The rest of the weight is split between ψ_h and ψ_l, and their relative weights depend on which accuracy level leads to a more pessimistic Bayesian posterior *given the realized message*. Intuitively, an agent who is averse to uncertainty in information accuracy ($\alpha > 0.5$) worries that good news' accuracy is low but bad news has high accuracy. An extreme form of pessimism can occur if $(1-\alpha)\psi_h + \alpha\psi_l < 50\%$. In this case, even the evaluation conditional on good news is (weakly) lower than the prior p.

The following proposition summarizes the predictions of Full Bayesian updating.

Proposition 1 *Suppose that an ε-α-maxmin agent uses Full Bayesian updating. In an uncertain information problem,*

1. *if $\varepsilon = 0$ and $\alpha = 0.5$, then her conditional evaluations coincide with the Bayesian evaluations conditional on simple information with accuracy level $\frac{\psi_h+\psi_l}{2}$;*

2. *as α increases, the conditional evaluations decrease;*

3. *as ε increases, the conditional evaluations become closer to p.*

4.1.2 Dynamically consistent updating

In an uncertain information problem, under Dynamically consistent updating (Hanany and Klibanoff, 2007), the evaluation of the bet conditional on message $m \in \{g, b\}$ is given by

$$u = Pr^{Bayes}(G|p, m, \max\{W(\psi_h, \psi_l; \varepsilon, \alpha), 50\%\}).$$

Unlike Full Bayesian updating, under Dynamically consistent updating, the as-if subjective information accuracy is the same regardless of the realized message. Specifically, the weight on 50% is always ε and the share of the rest of the weight assigned to ψ_l is α so long as the as-if information accuracy is still greater than 50%.

The interpretation of this formula is as follows. An agent who uses Dynamically consistent updating evaluates her contingent plan of choices *before the realization of information*. If the agent is averse to uncertainty in information accuracy ($\alpha > 0.5$), then she will prefer to under-react to information so that her ex-ante payoff is less dependent on the realization of that uncertainty.

Proposition 2 *Suppose an ε-α-maxmin agent uses Dynamically consistent updating. In an uncertain information problem,*

1. *if* $\varepsilon = 0$ *and* $\alpha = 0.5$*, then the conditional evaluations coincide with the Bayesian evaluations conditional on information with accuracy level* $\frac{\psi_h+\psi_l}{2}$*;*

2. *as either* ε *or* α *increases, the conditional evaluations become closer to* p*.*

4.1.3 Maximum likelihood updating

In an uncertain information problem, Maximum likelihood updating (Gilboa and Schmeidler, 1993) selects only the accuracy level(s) that is mostly likely given the realized message. Then the agent conducts Full Bayesian updating using the selected accuracy level(s). Since messages that confirm the prior are more likely to be accurate than not, agents update too much to them. By a similar logic, they update too little to messages that go against the prior. Formally, if $p \neq 50\%$, then the evaluation of the bet conditional on good news is given by

$$u = \begin{cases} Pr^{Bayes}(p, g, \psi_h), & \text{if } p > 50\% \\ Pr^{Bayes}(p, g, \psi_l), & \text{if } p < 50\% \end{cases}$$

and that conditional on bad news is

$$u = \begin{cases} Pr^{Bayes}(p, b, \psi_l), & \text{if } p > 50\% \\ Pr^{Bayes}(p, b, \psi_h), & \text{if } p < 50\% \end{cases}.$$

If $p = 50\%$, then the predictions of Maximum likelihood updating coincide with those of Full Bayesian updating.

The following proposition summarizes the properties of Maximum likelihood updating.

Proposition 3 *Suppose an* ε*-*α*-maxmin agent uses Maximum likelihood updating. In an uncertain information problem,*

1. *If* $p \neq 50\%$*, the conditional evaluations of the bet exhibit confirmation bias relative to those conditional on simple information with accuracy* $\frac{\psi_h+\psi_l}{2}$*. The measures of uncertainty attitudes,* ε *and* α*, do not affect the conditional evaluations.*

2. *If* $p = 50\%$*, conditional evaluations under Maximum likelihood updating coincide with those under Full Bayesian updating.*

4.1.4 Summary of theoretical implications

Consider an ε-α-maxmin agent whose attitudes toward uncertain information fall in the typical range: $\varepsilon > 0$ and $\alpha > 0.5$. Taking Bayesian learning with the corresponding simple information as

Theory	Aversion ($\alpha > 0.5$)	Insensitivity ($\varepsilon > 0$)
Full Bayesian updating	Pessimism	Under-reaction to news
Dynamically consistent updating	Under-reaction to news	Under-reaction to news
Maximum likelihood updating	$p \neq 50\%$: Confirmation bias (α and ε are irrelevant) $p = 50\%$: coincide with FBU	

Table 4.1: Summary of theoretical predictions in uncertain information problems

the benchmark, Table 4.1 summarizes the predictions of the updating rules I have discussed so far. The left panel of Figure 4.2 illustrates what the three main predictions, under-reaction, pessimism, and confirmation bias, each implies about the comparisons between belief updating with simple and uncertain information. Under-reaction and pessimism have the same directional predictions for good news but opposite for bad news. For neutral news, i.e. messages whose (midpoint) accuracy is 50%, under-reaction predicts that uncertain information accuracy does not have an effect on posteriors whereas pessimism predicts that posteriors conditional on uncertain information will be lower. The directional prediction of confirmation bias depends on the prior: the posteriors conditional on uncertain information are higher given high priors and lower given low priors.

If $\varepsilon = 0$ and $\alpha = 0.5$, then all theories except Maximum likelihood updating coincide with the benchmark.

4.2 Experimental results

Table 4.2 and the right panel of Figure 4.2 show the CEs of simple bets conditional on good news, bad news, and neutral news.[25] Additional statistical tests, including within- and between-subject t-tests, are in Table B.5. Perhaps the most salient empirical pattern is that the mean CEs conditional on compound and ambiguous good news are lower than the mean CE conditional on their simple counterpart for every combination of prior and (midpoint) information accuracy. This is consistent with the predictions of both under-reaction and pessimism, but not consistent with confirmation bias.

For bad news, the mean CEs conditional on compound and ambiguous news are higher compared to simple news in 3 out of 5 comparisons, and they are slightly lower or mixed in the other two. Since under-reaction and pessimism generate opposite direction predictions for bad news, the results can be explained by a combination of the two.

The mean CEs of a 70%-odds bet conditional on compound and ambiguous neutral news are significantly lower than that conditional on simple neutral news. For a 30%-odds bet, the mean CEs conditional on compound and ambiguous neutral news are statistically indistinguishable from

[25] Throughout the paper, I refer to the message "Red horse won" as good news for the Red bet and bad news for the Blue bet, and vice versa for the message "Blue horse won." In this section, I refer to any message whose (midpoint) accuracy is 50% as neutral news.

that conditional on simple neutral news. Again, the results are consistent with the combination of under-reaction and pessimism, but not with confirmation bias. Taken together, the empirical patterns resemble the prediction of Full Bayesian updating the most.

To further demonstrate the under-reaction and pessimism caused by uncertain information accuracy, for each uncertain information round I define *absolute pessimists/optimists* and *absolute under-/over-reactors*, two pairs of mutually exclusive categories, and then show that the former in each pair prevails.

First, I introduce some notations. Let $CE(p, m, \psi_h \text{ or } \psi_l)$ denote the conditional CE of a bet in an uncertain information round, where p is the prior of the bet, $m \in \{g, b\}$ indicates whether the realized message is *g*ood or *b*ad news for this bet, and the third argument represents the possible accuracy levels of the information.[26] Analogously, let $CE(p, m, \psi)$ be the conditional CE of a bet in a round with simple prior and simple information. Then, if the midpoint information accuracy in an uncertain information round is not 50%, define the *uncertainty premium* of a bet in this round as

$$Pm(p, m, \psi_h \text{ or } \psi_l) := CE(p, m, \frac{\psi_h + \psi_l}{2}) - CE(p, m, \psi_h \text{ or } \psi_l).$$

Note that $Pm(p, m, \psi_h \text{ or } \psi_l)$ may be missing for some subjects because calculating it requires that $CE(p, m, \frac{\psi_h+\psi_l}{2})$ be available in the data. In the rounds where the information accuracy is either 90% or 10%, since all news is neutral, I do not distinguish between good and bad news. The uncertainty premium of a bet in these rounds is defined as

$$Pm(p, -, 90\% \text{ or } 10\%) := CE(p, m', 50\%) - CE(p, m, 90\% \text{ or } 10\%),$$

where m and m' are the realized messages in the respective rounds.

Now I can define the categories, which are summarized in Table 4.3. A subject is an absolute pessimist in an uncertain information round if the uncertainty premiums of the two bets in this round are both weakly positive and at least one of them is strictly positive. Analogously, a subject is an absolute optimist if the two uncertainty premiums are both weakly negative and at least one is strictly negative.

In an uncertain information round where the midpoint information accuracy is not 50%, if a subject's uncertainty premium for the bet that receives good news is weakly positive, her uncertainty premium for the other bet is weakly negative, and at least one of the two is not zero, then I call this subject an absolute under-reactor. If, on the contrary, the bet that receives good news has a weakly negative uncertainty premium, the other one has a weakly positive premium, and at least one is not zero, then this subject is called an absolute over-reactor. In rounds where all news is neutral, I do not classify subjects into these two categories.

[26] I suppress notations for compound vs. ambiguous uncertainty when there is no risk of confusion.

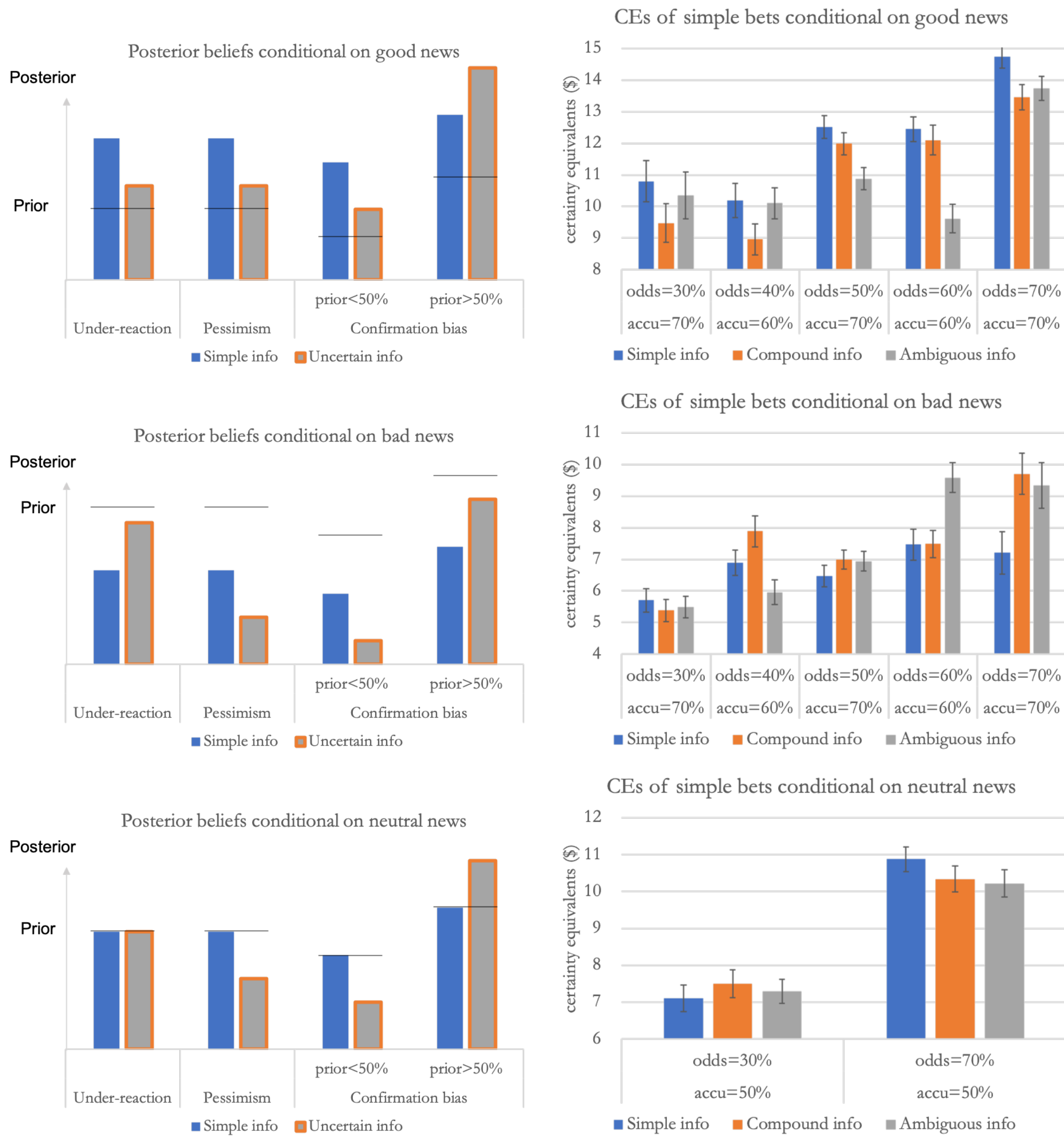

Figure 4.2: Simple priors with simple and uncertain information

Notes: The left panel of this figure illustrates what under-reaction, pessimism, and confirmation bias each predicts about the comparisons between belief updating with simple and uncertain information. (Neutral news refers to any message whose (midpoint) accuracy is 50%.) The right panel compares the mean CEs of simple bets conditional on simple, compound and ambiguous information in the experiment. Each group of bars correspond to a combination of prior and information. For example, "odds=30%, accu=70%" in the upper right graph represents tasks where the prior is 30% and the information is good news with 70% (midpoint) accuracy. Error bars represent 95% confidence intervals.

Prior	(Midpoint) Information accuracy	Good/Bad news	Type of information	Mean conditional CE	Standard error	N
30%	70%	good	simple	10.80	0.645	54
			compound	9.48	0.603	44
			ambiguous	10.35	0.739	47
40%	60%	good	simple	10.19	0.533	91
			compound	8.96	0.491	85
			ambiguous	10.10	0.490	60
50%	70%	good	simple	12.51	0.362	164
			compound	11.99	0.349	164
			ambiguous	10.88	0.346	165
60%	60%	good	simple	12.45	0.391	73
			compound	12.10	0.463	80
			ambiguous	9.61	0.452	105
70%	70%	good	simple	14.74	0.369	111
			compound	13.45	0.397	121
			ambiguous	13.74	0.381	118
30%	70%	bad	simple	5.70	0.368	111
			compound	5.38	0.355	121
			ambiguous	5.48	0.345	118
40%	60%	bad	simple	6.89	0.400	73
			compound	7.89	0.490	80
			ambiguous	5.95	0.390	105
50%	70%	bad	simple	6.47	0.345	165
			compound	6.99	0.306	163
			ambiguous	6.93	0.314	165
60%	60%	bad	simple	7.46	0.496	91
			compound	7.48	0.431	85
			ambiguous	9.58	0.474	60
70%	70%	bad	simple	7.20	0.672	54
			compound	9.70	0.651	44
			ambiguous	9.34	0.720	47
30%	50%		simple	7.10	0.361	163
			compound	7.50	0.374	164
			ambiguous	7.29	0.334	164
70%	50%		simple	10.88	0.336	163
			compound	10.34	0.349	164
			ambiguous	10.21	0.369	163

Table 4.2: Simple bets with additional information

Notes: This table compares the CEs of simple bets conditional on simple, compound and ambiguous information.

	Uncertainty premium	Bet that receives good news +	-
Bet that receives	+	Absolute pessimist	Absolute over-reactor
bad news	-	Absolute under-reactor	Absolute optimist

Table 4.3: Classification of subjects in an uncertain information round

Notes: This table summarizes the classification of subjects in an uncertain information round. To be classified into any of the four categories, the uncertainty premium of at least one bet in the round needs to be non-zero. For rounds whose midpoint information accuracy is 50%, I do not classify subjects as absolute over-/under-reactors.

Table 4.4 shows the percentages of each category in each round with uncertain information accuracy. In every round, there are more absolute under-reactors than absolute over-reactors. In all but one rounds, there are more absolute pessimists than absolute optimists. In aggregate, the percentage of absolute under-reactors is significantly higher for both compound and ambiguous information rounds. The aggregate percentage of absolute pessimists is also higher than that of absolute optimists, and the difference is significant for ambiguous information rounds.[27]

In summary, my experimental results show that uncertainty in the accuracy of information leads to under-reaction and pessimism. These two patterns are most consistent with the prediction of uncertainty-induced insensitivity and aversion together with Full Bayesian updating.

[27]In Appendix B.1, I consider two alternative categories: absolute confirmation bias and absolute contradiction bias. These two categories overlap with absolute over-/under-reactors, as an absolute over-reactor in a round with a confirmatory message is classified into the category of absolute confirmation bias. In all but one round, there are fewer absolute confirmation-biased subjects than absolute contradiction-biased subjects. This result together with the comparisons between mean CEs of bets suggests that uncertainty in information accuracy does not lead to prevalent confirmation bias.

Prior	Midpoint Information accuracy	Type of information	Absolute pessimists	Absolute optimists	p-value %(Abs. pess.) =%(Abs. opt.)	Absolute under-reactors	Absolute over-reactors	p-value %(Abs. under.) =%(Abs. over.)	N
(50%, 50%)	70%	Ambiguous	31.5%	19.4%	0.029	44.2%	18.2%	0	165
(60%, 40%)	60%	Ambiguous	31.0%	18.3%	0.128	43.7%	18.3%	0.007	71
(70%, 30%)	70%	Ambiguous	25.5%	23.4%	0.768	40.4%	19.1%	0.008	94
(70%, 30%)	50%	Ambiguous	25.5%	19.7%	0.31	-	-	-	137
Aggregate		Ambiguous	28.5%	20.1%	0.01	43.0%	18.5%	0	
(50%, 50%)	70%	Compound	21.2%	25.5%	0.425	43.0%	22.4%	0.001	165
(60%, 40%)	60%	Compound	27.4%	23.6%	0.586	36.8%	24.5%	0.107	106
(70%, 30%)	70%	Compound	27.6%	16.3%	0.057	39.8%	26.0%	0.059	123
(70%, 30%)	50%	Compound	27.7%	19.7%	0.172	-	-	-	137
Aggregate		Compound	25.6%	21.5%	0.164	40.4%	24.1%	0	

Table 4.4: Classification of subjects in each uncertain information round

Notes: This table shows the percentages of subjects that are classified into the four categories in each uncertain information round. Only subjects who face comparable belief updating problems in the uncertain information round and its corresponding simple information round are counted. In the rows under “Aggregate", I calculate the percentage of instances subjects are classified into each category, aggregated across the four or three rounds that are relevant for that category. The p-values are computed using Pearson’s chi-square goodness-of-fit tests.

5 Belief updating with uncertain priors

The belief updating rules studied in the previous section also make predictions on how people updating their beliefs when priors are uncertain. In this section, I briefly summarize these predictions and the experimental results on belief updating with uncertain priors and simple information. Details are relegated to Appendix A.

In an *uncertain prior* problem, the prior probability of G is either p_h or p_l with $p_l < p_h$, but the accuracy of additional information is known to be $\psi \geq 50\%$. Under both Full Bayesian updating and Dynamically consistent updating, an ε-α-maxmin agent's evaluation of the bet conditional on message $m \in \{g, b\}$ is

$$u = Pr^{Bayes}(G|W(p_h, p_l; \varepsilon, \alpha), m, \psi).$$

The uncertainty-induced insensitivity parameter ε is responsible for the degree of under-weighting of priors ("base-rate neglect"), and the uncertainty aversion parameter α corresponds to pessimism.

Under Maximum likelihood updating, the conditional evaluations are $u = Pr^{Bayes}(p_h, g, \psi)$ for good news and $u = Pr^{Bayes}(p_l, b, \psi)$ for bad news. This leads to over-reaction to news. For neutral news ($\psi = 50\%$), Maximum likelihood updating coincides with the other two updating rules.

The experimental result shows that posterior beliefs are more pessimistic if the priors are uncertain. This pattern is demonstrated both by comparing average conditional CEs of uncertain and simple bets and through subject classification. However, I do not find evidence for either under-weighting of priors or over-reaction to news. The overall result is in line with the prediction of Full Bayesian updating and Dynamically consistent updating coupled with uncertainty aversion.[28]

6 Structural analysis

In the previous sections, I show how different kinds of uncertainty directionally affect CEs of bets in a variety of comparisons. Next, I will integrate the comparisons using a representative-agent model and obtain quantitative measures of the effects of different kinds of uncertainty. To ensure that the structural estimates indeed capture the marginal effects of uncertainty, I need to adapt the previously-introduced theoretical framework so that the empirical model allows for nonstandard risk preferences and inherent belief updating biases that are unrelated to ambiguity and compound uncertainty.

[28]One potential reason for why there isn't evidence for under-weighting of priors when priors are uncertain is that even when priors are simple, base-rate neglect is already quite severe. (See Section 6.2.) As a result, there may not be enough room for additional base-rate neglect to be detected when priors become uncertain.

6.1 Risk preference

For risk preference, I use Prelec (1998)'s two-parameter function to model the CE of a bet with winning odds p and stake \$20:

$$CE(p) = M^{Prelec}(p) := \$20 \cdot exp(-b(-log(p))^a).$$

This model allows an agent to exhibit risk aversion/pessimism (b) and insensitivity (a) even when the decision problem does not involve compound uncertainty or ambiguity.[29] I assume that new information, compound uncertainty, and ambiguity do not affect an agent's risk attitudes (a and b), so they affect the CE of a bet only through their effects on the (subjective) winning odds of the bet. In other words, I assume that agents form a subjective winning odds and then apply the risk preference $M^{Prelec}(\cdot)$ to obtain the CE.

6.2 Inherent belief updating biases

Subjects' belief updating behaviors may deviate from the Bayes' rule for reasons unrelated to compound uncertainty and ambiguity. Figure B.1 and Table B.6 show the CEs of simple bets conditional on simple information and compare them to their Bayesian benchmarks whenever available.[30] For every task where the prior is not 50%, the mean conditional CE deviates from the Bayesian benchmark in the direction of under-weighting of priors ("base-rate neglect") (Kahneman and Tversky, 1973; Grether, 1980). This clear pattern demonstrates the importance of accounting for inherent belief updating biases when trying to identify the marginal effects of compound uncertain and ambiguity on belief updating.

To model belief updating with simple priors and simple information, I use the *generalized Bayes' rule* which allows for over- and under-weighting of priors, good news, and bad news (Möbius et al., 2014):

$$Pr^{GB}(G|p, g, \psi) = \frac{p^\beta \psi^{r_g}}{p^\beta \psi^{r_g} + (1-p)^\beta (1-\psi)^{r_g}},$$

$$Pr^{GB}(G|p, b, \psi) = \frac{p^\beta (1-\psi)^{r_b}}{p^\beta (1-\psi)^{r_b} + (1-p)^\beta \psi^{r_b}},$$

[29] Prelec (1998) uses the two-parameter function to model the probability weighting function in Prospect Theory. It is easy to show that the CE of a bet takes the same functional form if the agent uses Prelec's probability weighting function combined with a power utility function.

[30] For example, the Bayesian posterior belief given a prior of 50% and a good news with 70% accuracy is 70%. Therefore, the Bayesian benchmark for the CE of a 50%-odds bet conditional on 70%-accurate good news is the CE of a bet with an odds of 70%. The Bayesian posteriors updated from a 60% prior using a 60%-accurate good news is 69%. Hence I take the CE of a 70%-odds bet as its Bayesian benchmark.

Part	Prior	Information	Empirical model for CEs
1	Simple	No information	$CE(p) = M^{Prelec}(p)$
2	Simple	Simple	$CE(p, m, \psi) = M^{Prelec}\left(Pr^{GB}(G\|p, m, \psi)\right)$
3	Simple	Uncertain	$CE(p, g, \psi_h \text{ or } \psi_l) = M^{Prelec}\left(Pr^{GB}(G\|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha))\right)$ $CE(p, b, \psi_h \text{ or } \psi_l) = M^{Prelec}\left(Pr^{GB}(G\|p, b, W(\psi_l, \psi_h; \varepsilon, \alpha))\right)$
4	Uncertain	No information	$CE(p_h \text{ or } p_l) = M^{Prelec}\left(W(p_h, p_l; \varepsilon, \alpha)\right)$
5	Uncertain	Simple	$CE(p_h \text{ or } p_l, m, \psi) = M^{Prelec}\left(Pr^{GB}(G\|W(p_h, p_l; \varepsilon, \alpha), m, \psi)\right)$

Table 6.1: Summary of the empirical model

where β, r_g and r_b are non-negative numbers. The generalized Bayes' rule coincides with Bayes' rule when β, r_g and r_b all equal one.

6.3 The empirical model

The empirical model is based on the ε-α-maxmin preferences and Full Bayesian updating and is adapted to allow for nonstandard risk preferences and inherent belief updating biases. For the evaluation of uncertain bets in problems without belief updating, I use the ε-α-maxmin model to represent the subjective probability of winning:

$$Pr(p_h \text{ or } p_l) = W(p_h, p_l; \varepsilon, \alpha).$$

To model evaluations of simple bets conditional on uncertain information, I apply the ε-α-maxmin formula used in Full Bayesian updating to obtain the subjective information accuracy: $W(\psi_h, \psi_l; \varepsilon, \alpha)$ for good news and $W(\psi_l, \psi_h; \varepsilon, \alpha)$ for bad news. Then I apply this subjective information accuracy to the generalized Bayes' rule: $Pr^{GB}(G|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha))$ for good news and $Pr^{GB}(G|p, b, W(\psi_l, \psi_h; \varepsilon, \alpha))$ for bad news.[31]

To model evaluations of uncertain bets conditional on simple information, I apply the ε-α-maxmin formula used in both Full Bayesian updating and Dynamically consistent updating to obtain the subjective prior belief and then apply this subjective prior to the generalized Bayes' rule: $Pr^{GB}(G|W(p_h, p_l; \varepsilon, \alpha), m, \psi)$ for both good and bad news.

Note that because the generalized Bayes' rule shares the same monotonicity properties of the Bayes' rule, all the comparative statics results in Sections 3, 4 and 5 remain valid. Table 6.1 summarizes the empirical model for CEs of bets in each part of the experiment.

Importantly, I allow a subject to have different attitudes toward uncertain information and

[31]Although the empirical model is based on Full Bayesian updating, it can accommodate the qualitative prediction of Dynamically consistent updating, which is under-/over-reaction.

uncertain priors with or without belief updating. I also allow attitudes toward compound uncertainty to be different from attitudes toward ambiguity. In other words, I will separately estimate $3 \times 2 = 6$ potentially different ε's and also six α's.

The empirical model is identified using the experimental data. Specifically, the risk preference is identified by CEs of simple bets without new information. The parameters in the generalized Bayes' rule are identified by the differences between CEs of simple bets conditional on simple information and CEs of simple bets without new information. The six sets of ε's and α's are identified by the differences between CEs in tasks where the odds of bets or information accuracy is uncertain and those in the corresponding simple tasks.

6.4 Results

I estimate the empirical model using nonlinear least squares at the aggregate level assuming homogeneous parameters across subjects.[32] In the estimation, I do not impose constraints on the values of estimates.

Table 6.2 shows the estimates of the key parameters of interest. The numbers have simple interpretations. For example, the row corresponding to compound information accuracy states that the representative subject neglects 10% of the content of compound information. Also, she over-weights (under-weights) the high accuracy by 54% − 46% = 8% relative to the low accuracy when the news is bad (good). There are several salient patterns in Table 6.2. First, in both uncertain priors and uncertain information, the more pessimistic possibility receives a higher weight than the more optimistic one. This pattern holds for both compound uncertainty and ambiguity, which suggests that uncertainty of all types consistently leads to pessimism. Ambiguity tends to induce more pessimism than compound uncertainty, except that the results are mixed for uncertain priors in problems with belief updating. Second, the estimates of ε's are mostly positive and significant, and they tend to be larger for ambiguity than for compound uncertainty. One exception is ambiguity in priors in problems with belief updating, where the estimates are noisy. Last, the estimates of α are similar in magnitude across uncertainty in priors and uncertainty in information accuracy, and so are estimates of ε.[33]

Table 6.3 shows the estimates of the incidental parameters. The degree of risk aversion (b) is small and insignificant, but subjects exhibit strong and significant insensitivity (a) in their risk preference.

[32] Specifically, this procedure finds the parameters that minimize the sum of squared differences between the CE of a bet in the data and that generated by the empirical model. The sum is taken over tasks and subjects.

[33] The exceptions are that compound uncertainty in priors in problems with belief updating generates more pessimism, and that ambiguity in priors in problems with belief updating leads to more insensitivity.

Type of uncertainty		α	ε
Info accuracy	Compound	0.54 (0.02)	0.10 (0.05)
	Ambiguous	0.58 (0.02)	0.17 (0.05)
Priors (without updating)	Compound	0.55 (0.02)	0.14 (0.04)
	Ambiguous	0.61 (0.02)	0.25 (0.05)
Priors (with updating)	Compound	0.61 (0.03)	0.14 (0.08)
	Ambiguous	0.60 (0.03)	-0.08 (0.09)

Table 6.2: Aggregate estimates of α and ε

Notes: All parameters are assumed to be homogeneous among subjects. Numbers in parentheses are standard errors, which are computed by a bootstrap clustered at the subject level. The model is estimated using nonlinear least squares.

Parameter	Estimate (s.e.)
a	0.77 (0.04)
b	1.03 (0.03)
β	0.51 (0.04)
r_g	0.97 (0.07)
r_b	0.73 (0.06)

Table 6.3: Aggregate estimates of incidental parameters

Notes: All parameters are assumed to be homogeneous among subjects. Standard errors are computed by a bootstrap clustered at the subject level. The model is estimated using nonlinear least squares.

Regarding inherent belief updating biases unrelated to compound or ambiguous uncertainty, subjects under-weight priors by half. They also under-weight bad news by 27% while weighting good news in a Bayesian manner.

Overall, the structural estimation confirms the descriptive results.[34]

7 Individual-level relation between uncertainty attitudes for priors and information

In the last section, I show that the effects of uncertain priors and uncertain information have similar magnitudes for the representative subject. This result seems to suggest that attitudes towards uncertain priors and uncertain information manifest the same behavioral trait. However, drawing

[34]In Appendix B, I show results of the structural estimation at the individual level, with order effects, and with alternative models. The empirical patterns remain robust in these variants.

this conclusion would be premature as the similarity between different uncertainty attitudes at the aggregate level does not imply the same similarity at the individual level.

In this section, I investigate whether attitudes toward uncertainty in information accuracy and uncertainty in priors (with and without belief updating) are correlated at the individual level. If such correlations are strong and significant, then we can rather safely use knowledge about an agent's attitude toward one kind of uncertainty to make predictions about her attitudes toward the others. Otherwise, extrapolation is not warranted and we would need to study them separately.

Correlation analysis is challenging because different combinations of updating rules and uncertainty attitudes can generate similar behavior. Without knowing the updating rule a subject adheres to, it is sometimes difficult to pin down her uncertainty attitudes. For example, suppose that an ε-α-maxmin subject exhibits under-reaction to news but no pessimism in an uncertain information problem. Then this behavior is consistent with $\varepsilon > 0, \alpha = 0$ and Full Bayesian updating, but it is also consistent with $\varepsilon \geq 0, \alpha > 0$ and Dynamically consistent updating. To circumvent this identification issue, I focus attention on tests of correlation that are valid under ε-α-maxmin preferences and all three updating rules considered so far.[35] In Appendix G, I consider the full set of correlation tests that are valid under each model.

The correlation tests I construct are based on the signs of uncertainty premiums. For a bet whose prior is either p_h or p_l, define the sign of its uncertainty premium in a problem without belief updating as

$$SP(p_h \text{ or } p_l) = sign\left(CE(\frac{p_h + p_l}{2}) - CE(p_h \text{ or } p_l)\right) := \begin{cases} 1, & \text{if } CE(p_h \text{ or } p_l) < CE(\frac{p_h+p_l}{2}) \\ 0, & \text{if } CE(p_h \text{ or } p_l) = CE(\frac{p_h+p_l}{2}) \\ -1, & \text{if } CE(p_h \text{ or } p_l) > CE(\frac{p_h+p_l}{2}) \end{cases}.$$

For a simple bet with uncertain information, define the sign of uncertainty premium as

$$SP(p, m, \psi_h \text{ or } \psi_l) = sign\left(Pm(p, m, \psi_h \text{ or } \psi_l)\right),$$

where $Pm(\cdot, \cdot, \cdot \text{ or } \cdot)$ is defined in Section 4.2. Similarly, define the sign of uncertainty premium of an uncertain bet in a problem with belief updating as

$$SP(p_h \text{ or } p_l, m, \psi) = sign\left(Pm(p_h \text{ or } p_l, m, \psi)\right),$$

where $Pm(\cdot \text{ or } \cdot, \cdot, \cdot)$ is defined in Appendix A.2.

[35] These tests are also valid under many models considered in the Appendix.

The following proposition lays out the basis for the tests of correlations between different kinds of uncertainty attitudes.

Proposition 4 *Suppose that an ε-α-maxmin agent uses either Full Bayesian updating, Dynamically consistent updating, or Maximum likelihood updating and adapts it to the generalized Bayes' rule. Then*

1. *if the agent's attitudes toward uncertain information and uncertain priors (in problems without updating) are described by the same ε-α-maxmin preference, then*

$$SP(50\%, g, 90\% \textit{ or } 50\%) = SP(90\% \textit{ or } 50\%);$$

2. *if the agent's attitudes toward uncertain information and uncertain priors (in problems with updating) are described by the same ε-α-maxmin preference, then*

$$SP(50\%, g, 90\% \textit{ or } 50\%) = SP(90\% \textit{ or } 50\%, -, 50\%);$$

3. *if the agent's attitudes toward uncertain priors in problems with and without updating are described by the same ε-α-maxmin preference, then*

$$SP(90\% \textit{ or } 50\%, -, 50\%) = SP(90\% \textit{ or } 50\%) \;\; \textit{and} \;\; SP(10\% \textit{ or } 50\%, -, 50\%) = SP(10\% \textit{ or } 50\%).$$

To see why item 1 in the proposition is true, note that for an ε-α-maxmin agent who uses Full Bayesian updating adapted to the generalized Bayes' rule, the comparison between $CE(50\%, g, 90\%$ or $50\%)$ and $CE(50\%, g, 70\%)$ boils down to the comparison between $W(90\%, 50\%; \varepsilon, \alpha)$ and 70%. If the same ε and α apply to both uncertainty in information accuracy and uncertainty in priors (in problems without updating), then the same comparison between $W(90\%, 50\%; \varepsilon, \alpha)$ and 70% also determines the comparison between $CE(90\%$ or $50\%)$ and $CE(70\%)$. Moreover, this statement is also true if the agent uses the other two belief updating rules. This is because the conditional CEs under Dynamically consistent updating coincide with Full Bayesian updating for good news, and those under Maximum likelihood updating are the same as Full Bayesian updating if $p = 50\%$. Similar arguments also apply to items 2 and 3. In Appendix G, I show the proof of Proposition 4. In fact, the proposition also holds under several extensions of the smooth model and Segal's two-stage model. See Appendix E and F for more details.

I compute the correlation between the two sides of each equation in Proposition 4 to test for correlation between attitudes toward two different kinds of uncertainty. Table 7.1 shows the results.

Test	Correlation coefficient	
	Ambiguity	Compound
SP(50%, g, 90% or 50%)=SP(90% or 50%)	0 (0.99)	-0.08 (0.28)
SP(50%, g, 90% or 50%)=SP(90% or 50%,-,50%)	0.03 (0.71)	0 (0.97)
SP(90% or 50%,-,50%)=SP(90% or 50%)	0.26 (0)	0.15 (0.05)
SP(10% or 50%,-,50%)=SP(10% or 50%)	0.22 (0)	0.1 (0.19)

Table 7.1: Results of correlation tests

Notes: This table lists the correlation coefficients of the tests that are valid under Full Bayesian updating, Dynamically consistent updating and Maximum likelihood updating adapted to generalized Bayes' rule. Numbers in parentheses are p-values with the null hypothesis being that the correlation is zero.

While the correlations that involves attitudes toward uncertain information are all very close to zero, the correlations between attitudes toward uncertain priors with and without belief updating have larger magnitudes and, in most cases, high significance. Taken together, these results imply that with or without the presence of news, subjects have rather consistent attitudes toward uncertainty in the distribution over payoff-relevant states (“uncertainty in fundamentals"). In contrast, their attitudes toward uncertainty in information accuracy are distinct from how they treat uncertain fundamentals.

8 Individual-level relation between attitudes toward compound uncertainty and ambiguity

Structural estimates in Section 6 show that at the aggregate level, the insensitivity and uncertain aversion induced by ambiguity tend to have larger magnitudes than those induced by compound uncertainty. In this section, I examine the individual-level relation between attitudes toward compound uncertainty and ambiguity. On the one hand, compound uncertainty and ambiguity differ on whether the full probability distribution over states is explicitly specified. On the other hand, both types of uncertainty are more complex than simple risks. Hence, investigating the association between compound and ambiguity attitudes sheds light on the relative importance of “unknown unknown" and complexity in decisions under uncertainty.

If an agent treats compound and ambiguous information identically, then

$$CE^{Comp}(p, m, \psi_h \text{ or } \psi_l) = CE^{Amb}(p, m, \psi_h \text{ or } \psi_l)$$

for any prior p, message m and information accuracy ψ_h and ψ_l. Similar equations hold for uncertain priors with or without belief updating if an agent holds the same attitudes toward compound and ambiguous uncertainty in priors.

Among all cases where a subject's CE for a simple bet and its compound and ambiguous counterparts are all available, there are 39% where the CEs of the compound and ambiguous bets are identical. The analogous percentages for uncertain information and uncertain priors (in problems with updating) are 36% and 35%.[36] To construct benchmarks for these percentages where attitudes toward compound and ambiguous uncertainty are independent, I generate independent uniform random permutations of the compound CEs and ambiguous CEs among those that share the same corresponding simple CE.[37] Using the permuted data, I calculate the same three percentages as before. Among 500 simulations, the highest numbers are 22%, 23%, and 21% for uncertain priors (without updating), uncertain information, and uncertain priors (with updating), respectively. These numbers are significantly lower than the actual percentages of cases where a subject's compound CE is equal to her corresponding ambiguous CE, which implies that the match between compound and ambiguity attitudes is not merely coincidence. Also, I show in Table H.1 that the result is not simply driven by cases where the corresponding simple, compound, and ambiguous CEs are all the same, as the conclusion remains even if I exclude these cases. Moreover, there are more cases where the compound CE coincides with its corresponding ambiguous CE than where either of these two matches the simple CE. Taken together, my results show that compound uncertainty and ambiguity are often treated as the same by subjects.

9 Evidence from the stock market

In this section, I complement the experimental results with evidence from the US stock markets. Consistent with the lab findings, I show that stock prices react less sufficiently to analyst earnings forecasts with more uncertain accuracy. In addition, the decrease in reaction sufficiency occurs only for good news but not for bad news. These empirical patterns suggest that the experimental findings on learning from unknown information sources are externally valid and economically important.

Brokerage houses hire financial analysts to conduct research on publicly traded companies and to issue forecasts on their earnings. A large literature in accounting and finance has studied the

[36]If I do not require the simple CE to be available, the percentages are 39%, 35% and 35%, respectively.

[37]For example, there are 18 subjects who report $CE(60\%, g, 60\%) = 12$ and among these 18 subjects, there are 11 whose $CE^{Comp}(60\%, g, 90\% \text{ or } 30\%)$ is not missing. Hence, I randomly permute these 11 CEs which are conditional on compound information. Similarly, there are 13 subjects among the 18 whose $CE^{Amb}(60\%, g, 90\% \text{ or } 30\%)$ is not missing. I generate an independent random permutation of these 13 CEs conditional on ambiguous information.

information content of analysts' forecasts, the market reactions to them, and the factors that affect the sufficiency of these reactions (Kothari et al., 2016). What makes this setting suitable for studying unknown information sources is that forecasts differ in the uncertainty of their accuracy, depending on how familiar the issuing analysts are to the investors.

In Appendix I.1, I adapt the model with ε-α-maxmin preferences and Full Bayesian updating to a setting of stock investment. When the accuracy of an earnings forecast is uncertain, the reaction of the earnings expectation of an investor with typical uncertainty attitudes ($\varepsilon > 0$ and $\alpha > 0.5$) will be insufficient and biased downward. To the extent that stock price movement reflects the change in investors' earnings expectations, the stock price reactions to forecasts with uncertain accuracy will exhibit the same patterns.

To empirically test the theoretical predictions, I use data from three sources: quarterly earnings forecasts and earnings announcements from the Institutional Broker Estimate System (I/B/E/S) detail history file, stock returns from the Center for Research in Security Prices (CRSP), and firm characteristics from Compustat. I require stocks to be common shares (share codes 10 or 11) on the AMEX, NYSE, or NASDAQ (exchange codes 1, 2, or 3), and I exclude stocks with prices less than \$1 or market capitalization smaller than \$5 million. I restrict attention to earnings forecasts for quarters between January 1st, 1994 and June 30th, 2019,[38] but in order to construct analyst characteristics such as experience, I use data dated back to January 1st, 1984.

To measure the sufficiency of the market's reactions to analysts' forecasts, I calculate the correlation between the immediate price reactions and the price drifts that ensue, following the tradition in macro, finance, and accounting (e.g. Coibion and Gorodnichenko, 2015). Intuitively, if on average, immediate price reactions are followed by drifts in the same (opposite) direction, then the immediate reactions must be insufficient (excessive). The immediate price reaction to an earnings forecast is measured as the stock's size-adjusted returns[39] in the [-1,1]-trading day window centered on the forecast release, and the drift is the size-adjusted returns in the [2,64]-trading day period (which is roughly 3 months). To mitigate the confounds of other news events in the immediate reaction window, I only include observations where on the forecast announcement day, there is neither earnings announcement from the company nor earnings forecast announcements by any other analyst on the same company. Also, I restrict attention to forecast revisions, which can be naturally classified into good news and bad news. Following Gleason and Lee (2003), I define good news as a upward revision, which is a forecast that is higher than the issuing analyst's previous forecast on the same

[38] I do not include observations that date further back in time because the announcement dates recorded in I/B/E/S often differed from the actual dates by a couple of days prior to early 1990s.

[39] Size-adjusted returns are the stock's buy-hold returns minus the equal-weighted average returns of stocks in the same size decile in the same period.

quarterly earnings. Analogously, bad news is defined as a downward forecast revision. This leaves us with a final sample of 1,025,823 forecasts issued by 12,815 analysts on 10,712 stocks.

To proxy for the uncertainty of an analyst's forecast accuracy for a stock, I look at whether the analyst has a forecast record for that stock. At a point in time, an analyst has a forecast record for a stock if she has issued a quarterly earnings forecast on this stock before *and* the actual earnings of that quarter have been announced. This proxy is valid because prior research has shown that forecast accuracy is stock-specific and persistent (Park and Stice, 2000), that past forecast accuracy predicts future accuracy better than many other analyst attributes (Brown, 2001; Hilary and Hsu, 2013), and that investors learn about an analyst's forecast accuracy from her forecast record (Chen et al., 2005). I will henceforth refer to forecasts issued by analysts without (stock-specific) forecast records as "no-record forecasts" and the rest as "with-record forecasts."

In addition to the uncertainty in accuracy, no-record and with-record forecasts differ in other dimensions as well. Table I.3 provides summary statistics for a variety of characteristics of the forecasts, the issuing analysts, the covered stocks, and the information environment.[40] Variable definitions are in Table I.2. No-record forecasts on average have larger realized forecast errors. The companies they cover tend to be smaller, have higher and more volatile past returns and lower book-to-market ratios, and are followed by fewer analysts. The analysts without past records follow fewer stocks and industries. All variables in Table I.2 are included in the regressions.

Descriptive results show clear patterns that support the hypotheses. Figure 9.1 plots the average (size-adjusted) returns from one trading day before the forecast announcement to 1 trading day, 1 month, and 2 months after the forecast announcement, normalized by the average 3-month returns. For good news, the 1-day, 1-month and 2-month reactions are less sufficient for no-record forecasts. In contrast, for bad news, there is almost no difference in sufficiency between reactions to no-record and with-record forecasts. These results suggest that investors react insufficiently and pessimistically to no-record forecasts.[41]

The main specification of the regression analysis is as follows.

$$\begin{aligned}Ret[2,64]_i =& \eta_0 + \eta_1 Ret[-1,1]_i + \eta_2 NoRecord_i + \eta_3 GoodNews_i \\ &+ \eta_4 NoRecord_i \cdot GoodNews_i + \eta_5 Ret[-1,1]_i \cdot GoodNews_i + \eta_6 Ret[-1,1]_i \cdot NoRecord_i \\ &+ \eta_7 Ret[-1,1]_i \cdot NoRecord_i \cdot GoodNews_i + Controls_i + Controls_i \cdot Ret[-1,1]_i + TimeFE_i + \varepsilon_i .\end{aligned} \tag{1}$$

[40]Table I.4 provides summary statistics for all earnings forecasts issued between January 1st, 1994 and June 30th, 2019, including those that do not meet our data selection criteria.

[41]The summary statistics for unnormalized returns in windows with different lengths are in Table I.1.

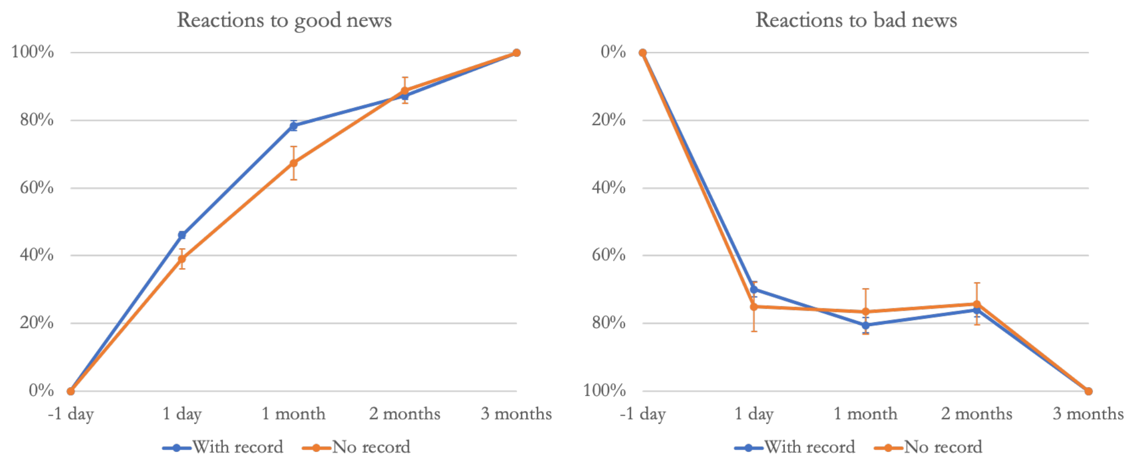

Figure 9.1: Reactions to forecast revisions

Notes: This figure shows the average size-adjusted returns from one trading day before the forecast announcement to 1 trading day, 1 month, and 2 months after the forecast announcement, normalized by the average 3-month returns. The left and right panels plot reactions to upward and downward forecast revisions, respectively. Error bars represent standard errors calculated using the delta method.

The dependent variable $Ret[2,64]_i$ is the size-adjusted stock returns in the [2,64]-trading day period after forecast i is announced, and $Ret[-1,1]_i$ is the immediate price reaction to forecast i. The indicator variable $NoRecord_i$ equals one if forecast i is a no-record forecast, and the variable equals zero otherwise. Following Gleason and Lee (2003), I define $GoodNews_i = 1$ if forecast i is an upward revision from the last forecast issued by the same analyst on the same stock's quarterly earnings, and $GoodNews_i = 0$ if the forecast is a downward revision. I include controls on the characteristics of the forecast, the issuing analyst, the covered stock, and the information environment (see Table I.2), as well as their interactions with $Ret[-1,1]$. I also include Year-Quarter dummies to control for unobserved time fixed effects on returns. In view of the descriptive results that immediate stock price reactions to no-record forecasts are less sufficient especially for good news, we expect the coefficient on the triple interaction, η_7, to be positive.

Table 9.1 shows the results from the regression analysis. Across the four specifications that differ on the set of controls and fixed effects, the coefficients on $NoRecord \times Ret[-1,1]$ and $NoRecord$ are small and insignificant, suggesting that for bad news, whether the issuing analyst of a forecast has a past record does not affect the sufficiency of immediate stock price reactions. In contrast, the coefficient on $Ret[-1,1] \times NoRecord \times GoodNews$ is consistently positive and significant. To interpret the magnitudes of the coefficients, the ratio between the price drift in the [2,64]-trading day window and the immediate reaction is larger for no-record good news than for with-record good news by around 10 percentage points. Taken together, the results imply that investors' reactions to earnings forecasts with more uncertain accuracy are more insufficient and pessimistic.

Dependent Var: Ret[2,64]	(1)	(2)	(3)	(4)
Ret[-1, 1]	0.0215	0.0173	0.343***	0.335**
	(0.0336)	(0.0333)	(0.100)	(0.100)
NoRecord	-0.000671	-0.00225	0.000814	0.000584
	(0.00287)	(0.00277)	(0.00213)	(0.00205)
NoRecord × Ret[-1, 1]	-0.0435	-0.0430	-0.0281	-0.0311
	(0.0626)	(0.0622)	(0.0474)	(0.0465)
GoodNews	0.0113***	0.0111***	0.0107***	0.0107***
	(0.00243)	(0.00211)	(0.00189)	(0.00177)
GoodNews × Ret[-1, 1]	0.0605†	0.0569	0.0480	0.0452
	(0.0351)	(0.0349)	(0.0294)	(0.0293)
NoRecord × GoodNews	0.00421	0.00440†	0.00102	0.00123
	(0.00269)	(0.00262)	(0.00253)	(0.00247)
NoRecord × GoodNews × Ret[-1, 1]	0.150*	0.150*	0.122†	0.124*
	(0.0624)	(0.0626)	(0.0626)	(0.0620)
Controls	N	N	Y	Y
Controls × Ret[-1,1]	N	N	Y	Y
Year-Quarter FE	N	Y	N	Y
Observations	1001418	1001417	894004	894004
R^2	0.001	0.010	0.004	0.014

Table 9.1: Sufficiency of stock market reactions to forecast revisions

Notes: This table reports the results of Regression (1). The dependent variable $Ret[2,64]$ is the size-adjusted stock returns in the [2,64]-trading day period after a forecast is announced, and $Ret[-1,1]$ is the immediate price reaction to a forecast. The variable $NoRecord$ indicates that a forecast is issued by an analyst with no stock-specific forecast record. The variable $GoodNews$ indicates an upward forecast revision. Control variables are characteristics of the forecast, the issuing analyst, the covered stock, and the information environment, summarized in Table I.2. Three-dimensional (stock, analyst, year-quarter) cluster-robust standard errors in parentheses. $\dagger p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001$

In Appendix I.3, I examine the robustness of the regression results. In Table I.5, I show that the signs of the coefficients are robust to changing the price drift window of the left-hand side variable in Specification (1). The effect sizes tend to increase as the drift window becomes longer, suggesting that the insufficiency of the immediate reactions are corrected gradually. Table I.6 shows the regression results for different cuts of the data. The results are robust when I only consider "high-innovation" forecast revisions, "isolated" forecasts, and forecasts announced after January 1st, 2004.[42] The main effect appears to be not solely driven by forecasts on small-cap stocks, as the

[42]Following Gleason and Lee (2003), a forecast revision is high-innovation if it falls outside the range

magnitude (though not the statistical significance) of the coefficient on the triple interaction term remains when I exclude all stocks with market capitalization smaller than \$2 billion. However, this coefficient vanishes if I only include large-cap stocks (market capitalization > \$10 billion), which may be due to the high concentration of sophisticated investors in those stocks. I also consider a specification that includes the interactions between Year-Quarter dummies and $Ret[-1, 1]$, and the results remain robust. Table I.7 reports the results of regressions that replace $Ret[-1, 1]$ and its interactions terms in Specification (1) with $Revision$ and its interactions terms. The variable $Revision$ is the difference between an analyst's revised forecast on earnings per share and the previous forecast, normalized by the stock price two trading days prior to the announcement of the revision. The results from this specification are similar: the price drift per unit of $Revision$ is larger for no-record good news than for with-record good news, but the difference is small and insignificant for bad news.

In sum, stock price reactions to earnings forecasts are less sufficient if they are issued by analysts with no forecast record. This phenomenon only happens for good news but not for bad news. These results corroborate the experimental finding that uncertainty in information accuracy leads to under-reaction to news and pessimism.

10 Conclusion

This paper studies the effects of uncertainty in information accuracy on belief updating using a controlled lab experiment and observational data from the stock market. In the experiment, a mean/mid-point preserving spread in the information accuracy leads subjects to react less to the information. Moreover, the reaction is biased toward the direction of bad news. The same two patterns also emerge in the stock market. I show that stock prices under-react more to earnings forecasts issued by analysts with no proven forecast record, and the under-reaction only occurs for good news but not for bad news. I examine the predictions of a wide variety of theories and find that a theory that combines Full Bayesian updating with uncertainty aversion and uncertainty-induced insensitivity best captures the empirical results.

between the issuing analyst's previous forecast and the previous consensus. (The consensus is the average of all forecasts available at the time.) High-innovation forecast revisions are likely to contain new information as they are not simply herding toward the consensus. "Isolated" forecasts are observations where in the 3-day window centered on the forecast announcement day, there is neither earnings announcement from the company nor forecast announcements by any other analysts on the same company. This filter further eliminates concerns that other news events might be driving $Ret[-1, 1]$. The focus on the period after 2004 is because a host of regulations on the financial analyst industry came into effect in 2002/2003 (Bradshaw et al., 2017), and the quality of forecast announcement time data in I/B/E/S improved after 2004 (Hirshleifer et al., 2019).

In the experiment, I compare the effects of uncertain information accuracy to those of uncertain priors. Both descriptive analysis and structural estimation show that uncertainty in priors leads to pessimism, and in problems without belief updating, it also induces insensitivity. Although the aggregate effects of uncertain information accuracy and uncertain priors are mostly similar in magnitudes, subjects' attitudes toward these two kinds of uncertainty are uncorrelated. The lack of correlation lends support to the view that uncertainty attitudes depend on the relevant issues as well as framing. Practically, it also suggests that knowing a person's attitude toward assets with unknown fundamentals does not help predict her reactions to information from unknown sources.

To separate the roles of complexity and incompletely-specified probabilities (“unknown unknowns") in generating belief updating biases, I compare the effects of compound and ambiguous information. Both kinds of information induce under-reaction and pessimism in the aggregate, but the effects of ambiguity are larger. At the individual level, there are many instances where a subject reacts to compound and ambiguous information in exactly the same way. Overall, the results suggest that both complexity and “unknown unknowns" play important roles in causing belief updating biases. Hence, to mitigate these biases, information needs to be both precise and simple to understand.

References

Abdellaoui, M., A. Baillon, L. Placido, and P. P. Wakker (2011) “The rich domain of uncertainty: Source functions and their experimental implementation,” *American Economic Review*, Vol. 101, No. 2, pp. 695–723.

Abdellaoui, M., P. Klibanoff, and L. Placido (2015) “Experiments on compound risk in relation to simple risk and to ambiguity,” *Management Science*, Vol. 61, No. 6, pp. 1306–1322.

Anantanasuwong, K., R. Kouwenberg, O. S. Mitchell, and K. Peijnenberg (2019) “Ambiguity attitudes about investments: Evidence from the field.”

Baillon, A., H. Bleichrodt, U. Keskin, O. l'Haridon, and C. Li (2017) “The effect of learning on ambiguity attitudes,” *Management Science*, Vol. 64, No. 5, pp. 2181–2198.

Baillon, A., Y. Halevy, and C. Li (2015) “Experimental elicitation of ambiguity attitude using the random incentive system,” *University of British Columbia working paper*.

Baillon, A., Z. Huang, A. Selim, and P. P. Wakker (2018) “Measuring Ambiguity Attitudes for All (Natural) Events,” *Econometrica*, Vol. 86, No. 5, pp. 1839–1858.

Barron, K. (2019) “Belief updating: Does the’good-news, bad-news’ asymmetry extend to purely financial domains?”Technical report, WZB Discussion Paper.

Becker, G. M., M. H. DeGroot, and J. Marschak (1964) “Measuring utility by a single-response sequential method,” *Behavioral science*, Vol. 9, No. 3, pp. 226–232.

Benjamin, D. J. (2019) “Errors in probabilistic reasoning and judgment biases,” in Bernheim, B. D., S. DellaVigna, and D. Laibson eds. *Handbook of Behavioral Economics - Foundations and Applications*, Vol. 2, pp. 69 – 186: North-Holland.

Bradshaw, M., Y. Ertimur, P. O’Brien et al. (2017) “Financial analysts and their contribution to well-functioning capital markets,” *Foundations and Trends® in Accounting*, Vol. 11, No. 3, pp. 119–191.

Brown, L. D. (2001) “How important is past analyst forecast accuracy?” *Financial Analysts Journal*, Vol. 57, No. 6, pp. 44–49.

Chateauneuf, A., J. Eichberger, and S. Grant (2007) “Choice under uncertainty with the best and worst in mind: Neo-additive capacities,” *Journal of Economic Theory*, Vol. 137, No. 1, pp. 538–567.

Chen, D. L., M. Schonger, and C. Wickens (2016) “oTree—An open-source platform for laboratory, online, and field experiments,” *Journal of Behavioral and Experimental Finance*, Vol. 9, pp. 88–97.

Chen, Q., J. Francis, and W. Jiang (2005) “Investor learning about analyst predictive ability,” *Journal of Accounting and Economics*, Vol. 39, No. 1, pp. 3–24.

Cheng, I.-H. and A. Hsiaw (2018) “Distrust in Experts and the Origins of Disagreement.”

Chew, S. H., B. Miao, and S. Zhong (2017) “Partial Ambiguity,” *Econometrica*, Vol. 85, No. 4, pp. 1239–1260.

Cohen, M., I. Gilboa, J. Jaffray, and D. Schmeidler (2000) “An experimental study of updating ambiguous beliefs,” *Risk, Decision and Policy*, Vol. 5, No. 2, p. 123–133.

Coibion, O. and Y. Gorodnichenko (2015) “Information rigidity and the expectations formation process: A simple framework and new facts,” *American Economic Review*, Vol. 105, No. 8, pp. 2644–78.

Corgnet, B., P. Kujal, and D. Porter (2012) "Reaction to public information in markets: how much does ambiguity matter?" *The Economic Journal*, Vol. 123, No. 569, pp. 699–737.

Coutts, A. (2019) "Good news and bad news are still news: Experimental evidence on belief updating," *Experimental Economics*, Vol. 22, No. 2, pp. 369–395.

De Filippis, R., A. Guarinno, P. Jehiel, and T. Kitagawa (2018) "Non-Bayesian updating in a social learning experiment."

Dempster, A. P. (1967) "Upper and lower probabilities induced by a multivalued mapping," *The Annals of Mathematical Statistics*, pp. 325–339.

Dimmock, S. G., R. Kouwenberg, and P. P. Wakker (2015) "Ambiguity attitudes in a large representative sample," *Management Science*, Vol. 62, No. 5, pp. 1363–1380.

Dominiak, A., P. Duersch, and J.-P. Lefort (2012) "A dynamic Ellsberg urn experiment," *Games and Economic Behavior*, Vol. 75, No. 2, pp. 625–638.

Edwards, W. (1968) "Conservatism in human information processing," in Kleinmuntz, B. and R. B. Cattell eds. *Formal Representation of Human Judgment*: John Wiley and Sons.

Eichberger, J., S. Grant, and D. Kelsey (2007) "Updating Choquet beliefs," *Journal of Mathematical Economics*, Vol. 43, No. 7, pp. 888 – 899.

Eil, D. and J. M. Rao (2011) "The good news-bad news effect: asymmetric processing of objective information about yourself," *American Economic Journal: Microeconomics*, Vol. 3, No. 2, pp. 114–38.

Ellsberg, D. (1961) "Risk, ambiguity, and the Savage axioms," *The quarterly journal of economics*, pp. 643–669.

——— (2015) *Risk, ambiguity and decision*: Routledge.

Enke, B. and T. Graeber (2019) "Cognitive Uncertainty."

Epstein, L. G. and M. Schneider (2003) "Recursive multiple-priors," *Journal of Economic Theory*, Vol. 113, No. 1, pp. 1–31.

——— (2008) "Ambiguity, information quality, and asset pricing," *The Journal of Finance*, Vol. 63, No. 1, pp. 197–228.

Epstein, L. and Y. Halevy (2019) “Hard-to-interpret signals.”

Ert, E. and S. T. Trautmann (2014) “Sampling experience reverses preferences for ambiguity,” *Journal of Risk and Uncertainty*, Vol. 49, No. 1, pp. 31–42.

Fox, C. R. and A. Tversky (1995) “Ambiguity aversion and comparative ignorance,” *The Quarterly Journal of Economics*, Vol. 110, No. 3, pp. 585–603.

Fryer, R. G., P. Harms, and M. O. Jackson (2018) “Updating beliefs when evidence is open to interpretation: Implications for bias and polarization,” *Journal of the European Economic Association*.

Gilboa, I. and M. Marinacci (2016) “Ambiguity and the Bayesian paradigm,” in *Readings in Formal Epistemology*, pp. 385–439: Springer.

Gilboa, I. and D. Schmeidler (1989) “Maxmin expected utility with non-unique prior,” *Journal of Mathematical Economics*, Vol. 18, No. 2, pp. 141–153.

——— (1993) “Updating ambiguous beliefs,” *Journal of economic theory*, Vol. 59, No. 1, pp. 33–49.

Gillen, B., E. Snowberg, and L. Yariv (forthcoming) “Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study,” *Journal of Political Economy*.

Gleason, C. A. and C. M. Lee (2003) “Analyst forecast revisions and market price discovery,” *The Accounting Review*, Vol. 78, No. 1, pp. 193–225.

Greiner, B. (2015) “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, Vol. 1, No. 1, pp. 114–125.

Grether, D. M. (1980) “Bayes rule as a descriptive model: The representativeness heuristic,” *The Quarterly Journal of Economics*, Vol. 95, No. 3, pp. 537–557.

Gul, F. and W. Pesendorfer (2014) “Expected uncertain utility theory,” *Econometrica*, Vol. 82, No. 1, pp. 1–39.

——— (2018) “Evaluating Ambiguous Random Variables and Updating by Proxy.”

Halevy, Y. (2007) “Ellsberg revisited: An experimental study,” *Econometrica*, Vol. 75, No. 2, pp. 503–536.

Hanany, E. and P. Klibanoff (2007) "Updating preferences with multiple priors," *Theoretical Economics*, Vol. 2, No. 3, pp. 261–298.

Hansen, L. and T. J. Sargent (2001) "Robust control and model uncertainty," *American Economic Review*, Vol. 91, No. 2, pp. 60–66.

Heath, C. and A. Tversky (1991) "Preference and belief: Ambiguity and competence in choice under uncertainty," *Journal of risk and uncertainty*, Vol. 4, No. 1, pp. 5–28.

Hilary, G. and C. Hsu (2013) "Analyst forecast consistency," *The Journal of Finance*, Vol. 68, No. 1, pp. 271–297.

Hirshleifer, D., Y. Levi, B. Lourie, and S. H. Teoh (2019) "Decision fatigue and heuristic analyst forecasts," *Journal of Financial Economics*, Vol. 133, No. 1, pp. 83–98.

Jaffray, J.-Y. (1992) "Bayesian updating and belief functions," *IEEE transactions on systems, man, and cybernetics*, Vol. 22, No. 5, pp. 1144–1152.

Kahneman, D. and A. Tversky (1973) "On the psychology of prediction.," *Psychological review*, Vol. 80, No. 4, p. 237.

——— (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Vol. 47, No. 2, pp. 263–292.

Kala, N. (2017) "Learning, adaptation and climate uncertainty: Evidence from Indian agriculture."

Keynes, J. M. (1921) "A Treatise on Probability."

Klibanoff, P., M. Marinacci, and S. Mukerji (2005) "A smooth model of decision making under ambiguity," *Econometrica*, Vol. 73, No. 6, pp. 1849–1892.

——— (2009) "Recursive smooth ambiguity preferences," *Journal of Economic Theory*, Vol. 144, No. 3, pp. 930–976.

Knight, F. (1921) "Risk, Uncertainty and Profit."

Kothari, S., E. So, and R. Verdi (2016) "Analysts' forecasts and asset pricing: A survey," *Annual Review of Financial Economics*, Vol. 8, pp. 197–219.

Machina, M. J. and M. Siniscalchi (2014) "Ambiguity and ambiguity aversion," in *Handbook of the Economics of Risk and Uncertainty*, Vol. 1, pp. 729–807: Elsevier.

Mikhail, M. B., B. R. Walther, and R. H. Willis (1997) "Do security analysts improve their performance with experience?" *Journal of Accounting Research*, Vol. 35, pp. 131–157.

Möbius, M. M., M. Niederle, P. Niehaus, and T. S. Rosenblat (2014) "Managing Self-Confidence."

Moreno, O. M. and Y. Rosokha (2016) "Learning under compound risk vs. learning under ambiguity–an experiment," *Journal of Risk and Uncertainty*, Vol. 53, No. 2-3, pp. 137–162.

Ngangoué, K. (2018) "Learning under Ambiguity: An Experiment on Gradual Information Processing."

Olszewski, W. (2007) "Preferences over sets of lotteries," *The Review of Economic Studies*, Vol. 74, No. 2, pp. 567–595.

Park, C. W. and E. K. Stice (2000) "Analyst forecasting ability and the stock price reaction to forecast revisions," *Review of Accounting Studies*, Vol. 5, No. 3, pp. 259–272.

Pires, C. P. (2002) "A rule for updating ambiguous beliefs," *Theory and Decision*, Vol. 53, No. 2, pp. 137–152.

Prelec, D. (1998) "The probability weighting function," *Econometrica*, pp. 497–527.

Schmeidler, D. (1989) "Subjective probability and expected utility without additivity," *Econometrica: Journal of the Econometric Society*, pp. 571–587.

Segal, U. (1987) "The Ellsberg paradox and risk aversion: An anticipated utility approach," *International Economic Review*, pp. 175–202.

——— (1990) "Two-stage lotteries without the reduction axiom," *Econometrica: Journal of the Econometric Society*, pp. 349–377.

Shafer, G. (1976) *A mathematical theory of evidence*, Vol. 1: Princeton university press Princeton.

Shishkin, D. ⓡ P. Ortoleva (2019) "Ambiguous information and dilation: an experiment."

Trautmann, S. and G. van de Kuilen (2015) "Ambiguity attitudes," in Keren, G. and G. Wu eds. *The Wiley Blackwell Handbook of Judgment and Decision Making*, pp. 89–116: John Wiley Sons Ltd.

Zhang, X. F. (2006) "Information uncertainty and stock returns," *The Journal of Finance*, Vol. 61, No. 1, pp. 105–137.

A Details on belief updating with uncertain priors

In this section, I present the theories and experimental evidence on belief updating with uncertain priors in detail.

A.1 Theories

In an *uncertain prior* problem, the prior probability of G is either p_h or p_l with $p_l < p_h$, but the accuracy of additional information is known to be $\psi \geq 0.5$. As in the previous section, I will apply several belief updating rules to the ε-α-maxmin model and compare their predictions in uncertain prior problems. Proofs of results in this subsection are in Appendix C.2.

A.1.1 Full Bayesian updating

Under Full Bayesian updating, an ε-α-maxmin agent behaves as if the prior is a weighted average of p_h, p_l and 50%, and updates accordingly in a Bayesian manner. The weight ε is always applied to 50%, and the rest of the weight, $1 - \varepsilon$, is split between p_h and p_l. Since regardless of the realized message, p_l always leads to a more pessimistic posterior than p_h, the former will receive α proportion of the rest of the weight. Hence, the Full Bayesian evaluation conditional on message $m \in \{g, b\}$ is

$$u = Pr^{Bayes}(G|W(p_h, p_l; \varepsilon, \alpha), m, \psi).$$

The insensitivity parameter ε is responsible for the degree of under-weighting of priors in belief updating, and the aversion parameter α corresponds to pessimism.

The following proposition summarizes the predictions of Full Bayesian updating in uncertain prior problems.

Proposition 5 *Suppose that an ε-α-maxmin agent uses Full Bayesian updating. In an uncertain prior problem,*

1. *if $\varepsilon = 0$ and $\alpha = 0.5$, then her conditional evaluations coincide with the Bayesian conditional evaluations given a simple prior $\frac{p_h+p_l}{2}$;*
2. *as α increases, the conditional evaluations decrease;*
3. *as ε increases, the evaluation conditional on good news becomes closer to ψ and that conditional on bad news becomes closer to $1 - \psi$.*

A.1.2 Dynamically consistent updating

In an uncertain prior problem, the conditional evaluations under Dynamically consistent updating are the same as those under Full Bayesian updating. Under Dynamically consistent updating, an agent who is averse to uncertainty ($\alpha > 0.5$) prefers to make choices so that her ex-ante payoff is less dependent on the realization of that uncertainty. When the uncertainty is in priors, mitigating ex-ante payoff exposure to uncertainty requires refraining from taking the bet. This coincides with Full Bayesian updating under which an uncertainty averse agent tries to mitigate ex-post payoff exposure to uncertainty. The following proposition summarizes the results.

Proposition 6 *In an uncertain prior problem, Dynamically consistent updating has the same predictions as Full Bayesian updating for an ε-α-maxmin agent.*

A.1.3 Maximum likelihood updating

In uncertain prior problems, the prior(s) that is most likely to generate the realized message is selected and updated under Maximum likelihood updating. Since good news is more likely to be generated from high priors and bad news from low priors, agents will over-react to news. Formally, if $\psi \neq 50\%$, then the evaluation of the bet conditional on good news is given by

$$u = Pr^{Bayes}(p_h, g, \psi)$$

and that conditional on bad news is

$$u = Pr^{Bayes}(p_l, b, \psi).$$

If $\psi = 50\%$, then the conditional evaluations coincide with Full Bayesian updating.

The following proposition summarizes the properties of Maximum likelihood updating.

Proposition 7 *Suppose an ε-α-maxmin agent uses Maximum likelihood updating. In an uncertain prior problem,*

1. *If $\psi \neq 50\%$, the conditional evaluations of the bet exhibit over-reaction relative to those given the simple prior $\frac{p_h+p_l}{2}$. The measures of uncertainty attitudes, ε and α, do not affect the conditional evaluations.*

2. *If $\psi = 50\%5$, conditional evaluations under Maximum likelihood updating coincide with those under Full Bayesian updating.*

<table>
<tr><th>Theory</th><th>Aversion ($\alpha > 0.5$)</th><th>Insensitivity ($\varepsilon > 0$)</th></tr>
<tr><td>Full Bayesian updating &
Dynamically consistent updating</td><td>Pessimism</td><td>Under-weighting of priors</td></tr>
<tr><td>Maximum likelihood updating</td><td colspan="2">$\psi \neq 50\%$: Over-reaction to news (α and ε are irrelevant)
$\psi = 50\%$: coincide with FBU</td></tr>
</table>

Table A.1: Summary of theoretical predictions in uncertain prior problems

A.1.4 Summary of theoretical implications

Consider an ε-α-maxmin agent whose attitudes toward uncertain priors fall in the typical range: $\varepsilon > 0$ and $\alpha > 0.5$. Taking Bayesian learning with the corresponding simple prior as the benchmark, Table A.1 summarizes the predictions of the three updating rules I have discussed so far. The left panel of Figure A.1 illustrates what the three main predictions, under-weighting of priors, pessimism, and over-reaction to news, each implies about the comparisons between belief updating with simple and uncertain priors.

If $\varepsilon = 0$ and $\alpha = 0.5$, then all theories except Maximum likelihood updating coincide with the benchmark.

A.2 Experimental results

Table A.2 and the right panel of Figure A.1 show the CEs of simple, compound and ambiguous bets conditional on simple information. Additional statistical tests, including within- and between-subject t-tests, are in Table A.5. Among the twelve combinations of prior and information accuracy, the mean conditional CE given the compound prior is lower than its simple counterpart in 8 comparisons, and the mean conditional CE given the ambiguous prior is lower in 7 comparisons. This suggests, though not strongly, that uncertain priors lead to pessimism in the conditional CEs. There is no clear pattern of either under-weighting of priors or over-reaction to news.

Similar as in the comparison between simple information and uncertain information, I define *absolute pessimists/optimists* and *absolute prior under-/over-weighters* for each uncertain prior round, and then compare their relative prevalence.

In an uncertain prior round, if the prior of a bet might either be p_h or p_l, the realized message is $m \in \{g, b\}$, and the information accuracy ψ is not 50%, then define the uncertainty premium of this bet in this round as

$$Pm(p_h \text{ or } p_l, m, \psi) := CE(\frac{p_h + p_l}{2}, m, \psi) - CE(p_h \text{ or } p_l, m, \psi).$$

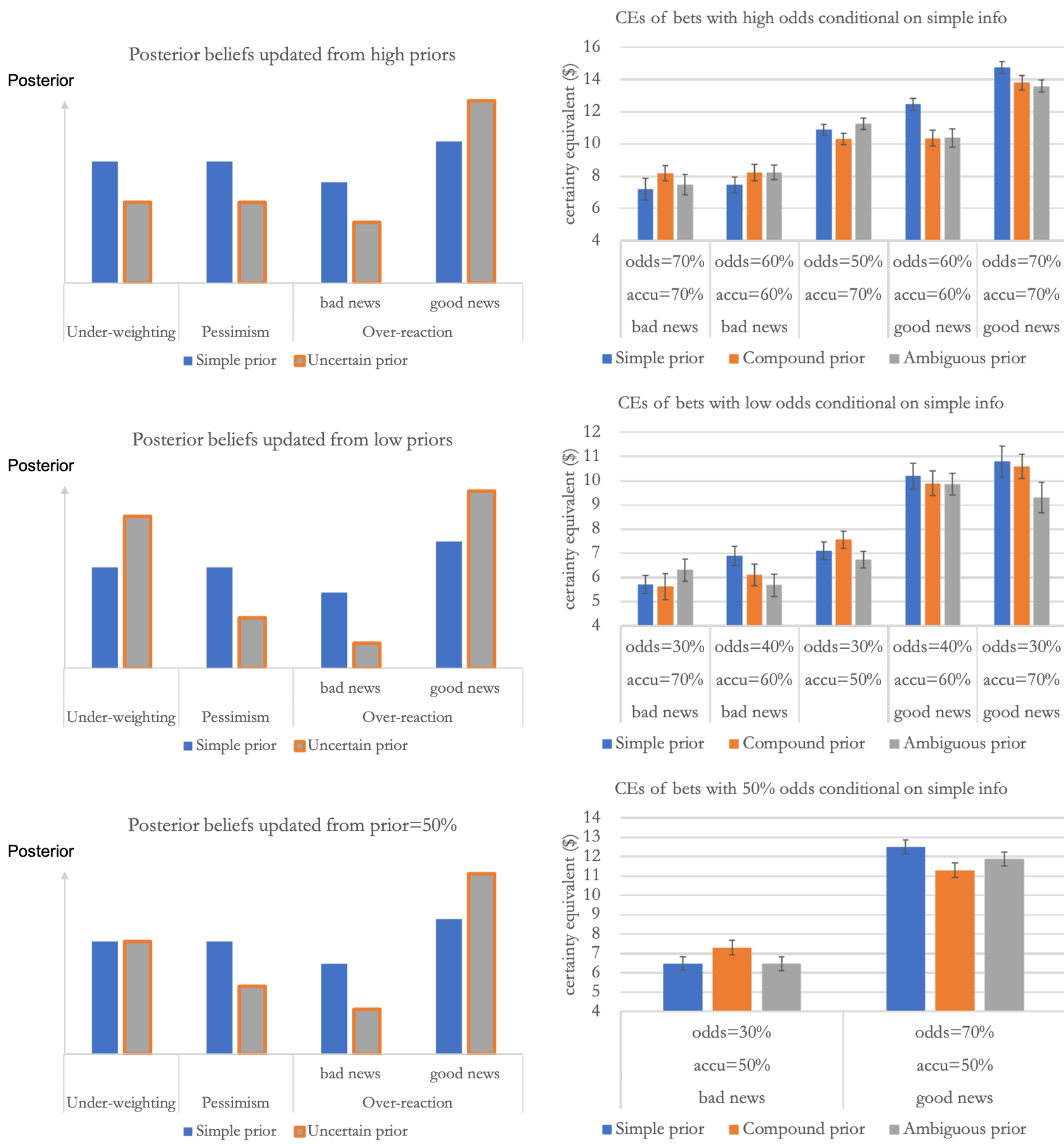

Figure A.1: Simple and uncertain priors with simple information

Notes: The left panel of this figure illustrates what under-weighting of priors, pessimism, and over-reaction to news each predicts about the comparisons between belief updating with simple and uncertain priors. (High, low, and medium priors refer to priors that are higher, lower, and equal to 50%.) The right panel compares the mean CEs of simple, compound, and ambiguous bets conditional on simple information in the experiment. Each group of bars correspond to a combination of prior and information. For example, “odds=70%, accu=70%, bad news” in the upper right graph represents tasks where the (midpoint) prior is 70% and the information is bad news with 70% accuracy. Error bars represent +/- one standard error.

(Midpoint) Prior	Information accuracy	Good/Bad news	Type of prior	Mean conditional CE	Standard error	N
30%	70%	bad	simple	5.70	0.368	111
			compound	5.62	0.529	71
			ambiguous	6.30	0.449	106
40%	60%	bad	simple	6.89	0.400	73
			compound	6.10	0.443	89
			ambiguous	5.68	0.464	74
30%	50%		simple	7.10	0.361	163
			compound	7.56	0.349	164
			ambiguous	6.74	0.341	165
40%	60%	good	simple	10.19	0.533	91
			compound	9.89	0.502	76
			ambiguous	9.85	0.446	89
30%	70%	good	simple	10.80	0.645	54
			compound	10.59	0.501	94
			ambiguous	9.31	0.642	59
70%	70%	bad	simple	7.20	0.672	54
			compound	8.17	0.472	94
			ambiguous	7.47	0.617	59
60%	60%	bad	simple	7.46	0.496	91
			compound	8.21	0.499	76
			ambiguous	8.22	0.445	89
70%	50%		simple	10.88	0.336	163
			compound	10.31	0.363	164
			ambiguous	11.27	0.353	165
60%	60%	good	simple	12.45	0.391	73
			compound	10.36	0.480	89
			ambiguous	10.37	0.566	75
70%	70%	good	simple	14.74	0.369	111
			compound	13.79	0.465	71
			ambiguous	13.58	0.375	106
50%	70%	bad	simple	6.47	0.345	165
			compound	7.29	0.385	164
			ambiguous	6.46	0.354	164
50%	70%	good	simple	12.51	0.362	164
			compound	11.30	0.379	165
			ambiguous	11.88	0.351	164

Table A.2: Bets with simple information

Notes: This table compares the CEs of simple, compound and ambiguous bets, all conditional on simple information.

	Uncertainty premium	Red bet +	Red bet -
Blue bet	+	Absolute pessimist	Absolute prior over-weighter
	-	Absolute prior under-weighter	Absolute optimist

Table A.3: Classification of subjects in an uncertain prior round

Notes: This table summarizes the classification of subjects in an uncertain prior round. To be classified into any of the four categories, the uncertainty premium of at least one bet in the round needs to be non-zero. For rounds whose midpoint prior is (50%, 50%), I do not classify subjects as absolute prior under-/over-weighters.

If $\psi = 50\%$, then I define the uncertainty premium as

$$Pm(p_h \text{ or } p_l, -, 50\%) := CE(\frac{p_h + p_l}{2}, m', 50\%) - CE(p_h \text{ or } p_l, m, 50\%),$$

where m and m' are the realized messages in the respective rounds.

The classification of subjects is summarized in Table A.3. Same as in an uncertain information round, a subject is classified as an absolute pessimist in an uncertain prior round if the uncertainty premiums for the two bets in this round are both weakly positive and at least one of them is strictly positive. An absolute optimist, on the other hand, is a subject whose uncertainty premiums for the two bets in this round are both weakly negative but not both zero.

In uncertain prior rounds where the midpoint prior is not (50%, 50%), a subject is classified as an absolute prior under-weighter if the uncertainty premium of the Red bet is weakly positive, that of the Blue bet is weakly negative, and one of the two is not zero.[43] Analogously, a subject is a prior over-weighter if the uncertainty premium of the Red bet is weakly negative, that of the Blue bet is weakly positive, and one of the two is not zero. I do not classify prior under-/over-weighters for rounds where the midpoint prior is (50%, 50%).

Table A.4 shows the percentages of each of the four categories in all eight rounds with uncertain priors and simple information. In all rounds but one, there are more absolute pessimists than optimists, and the percentage of the former aggregated across rounds is also significantly higher than the latter for both compound and ambiguous prior rounds. This further confirms that uncertain priors lead to pessimism. In contrast, there isn't strong evidence for either the under-weighting or over-weighting of priors. In three out of six rounds, there are more absolute prior under-weighters than over-weighters; in the other three rounds, the opposite is true.

[43]Recall that the Red bet always has a (midpoint) prior weakly higher than 50%.

Taken together, my experimental results suggest that in problems with belief updating, uncertainty in priors leads to pessimism. This pattern is consistent with the combination of uncertainty aversion and either Full Bayesian updating or Dynamically consistent updating. Under-weighting of priors, which is the prediction of uncertainty-induced insensitivity together with these two updating rules, is not borne out in the data.

Midpoint prior	Information accuracy	Type of prior	Absolute pessimists	Absolute optimists	p-value %(Abs. pess.) =%(Abs. opt.)	Absolute prior under-weighters	Absolute prior over-weighters	p-value %(Abs. negl.) =%(Abs. over.)	N
(50%, 50%)	70%	Amb	31.1%	21.3%	0.084	-	-	-	164
(60%, 40%)	60%	Amb	32.5%	24.7%	0.366	27.3%	33.8%	0.466	77
(70%, 30%)	70%	Amb	29.0%	16.7%	0.032	41.3%	31.9%	0.196	138
(70%, 30%)	50%	Amb	27.0%	18.4%	0.104	29.4%	35.6%	0.331	163
Aggregate		Amb	29.5%	18.8%	0.001	33.3%	33.9%	0.9	
(50%, 50%)	70%	Comp	32.3%	22.0%	0.072	-	-	-	164
(60%, 40%)	60%	Comp	39.3%	15.4%	0	35.9%	26.5%	0.198	117
(70%, 30%)	70%	Comp	21.3%	25.5%	0.67	31.9%	40.4%	0.493	47
(70%, 30%)	50%	Comp	25.2%	22.1%	0.569	33.1%	30.7%	0.695	163
Aggregate		Comp	30.5%	20.8%	0.002	33.9%	30.6%	0.449	

Table A.4: Classification of subjects in each uncertain prior round

Notes: This table shows the percentages of subjects that are classified into the four categories for each uncertain prior round. Only subjects who face comparable belief updating problems in the uncertain prior round and its corresponding simple prior round are counted. In the rows under “Aggregate", I calculate the percentage of instances subjects are classified into each category, aggregated across the four or three rounds that are relevant for that category. The p-values are computed using Pearson’s chi-square goodness-of-fit tests.

Prior and info	Type of prior	within-subject $\overline{CE}(simp) - \overline{CE}(unc)$	N	between-subject $\overline{CE}(simp) - \overline{CE}(unc)$	N(simp)	N(unc)
odds=30%, accu=70%	compound	-0.47 (0.52)	32	0.01 (0.99)	111	39
bad news	ambiguous	-0.21 (0.58)	95	-2.39 (0.06)	111	11
odds=40%, accu=60%	compound	-0.11 (0.84)	57	1.8 (0.03)	73	32
bad news	ambiguous	0.9 (0.26)	30	1.28 (0.08)	73	44
odds=30%, accu=50%	compound	-0.45 (0.18)	163			
neutral news	ambiguous	0.42 (0.16)	163			
odds=40%, accu=60%	compound	0.63 (0.29)	60	-0.75 (0.57)	91	16
good news	ambiguous	0.34 (0.61)	47	0.33 (0.71)	91	42
odds=30%, accu=70%	compound	0.13 (0.93)	15	0.09 (0.92)	54	79
good news	ambiguous	1.33 (0.06)	43	1.11 (0.43)	54	16
odds=70%, accu=70%	compound	-1.73 (0.02)	15	-0.78 (0.34)	54	79
bad news	ambiguous	0.12 (0.85)	43	-1.61 (0.25)	54	16
odds=60%, accu=60%	compound	-0.1 (0.86)	60	-0.29 (0.82)	91	16
bad news	ambiguous	0.02 (0.97)	47	-0.4 (0.63)	91	42
odds=70%, accu=50%	compound	0.56 (0.12)	163			
neutral news	ambiguous	-0.34 (0.27)	163			
odds=60%, accu=60%	compound	1.96 (0)	57	2.11 (0.02)	73	32
good news	ambiguous	0.97 (0.2)	31	2.47 (0)	73	44
odds=70%, accu=70%	compound	-0.03 (0.96)	32	1.74 (0.02)	111	39
good news	ambiguous	0.96 (0.01)	95	2.01 (0.1)	111	11
odds=50%, accu=70%	compound	-0.88 (0.03)	164			
bad news	ambiguous	-0.04 (0.9)	164			
odds=50%, accu=70%	compound	1.26 (0)	164			
good news	ambiguous	0.63 (0.1)	164			

Table A.5: Comparison between CEs of uncertain and simple bets conditional on simple information, with and without anchoring

Notes: This table shows the differences in mean conditional CEs between uncertain prior problems and simple prior problems. Numbers in parentheses are p-values in t-tests, and N is the number of subjects included. For example, the top row of the table states that there are 32 subjects who receive bad news both in the compound information problem and in the simple information problem where the prior is 30% and (midpoint) information accuracy 70%. Among these subjects, the difference in mean conditional CEs between these two problems is -$0.47 and the p-value of the paired t-test is 0.52. Thirty-nine subjects receive bad news in the compound prior problem but not in the simple problem, and there are 111 subjects who receive bad news in the simple problem in total. The difference between the mean conditional CE of the simple bet of the latter group and that of the compound bet of the former group is 0.01, and the p-value of the unpaired t-test is 0.99.

Prior	Midpoint Information accuracy	Type of information	Absolute confirmation bias	Absolute contradiction bias	N
(60%, 40%)	60%	Ambiguous	21	23	71
(70%, 30%)	70%	Ambiguous	23	33	94
(70%, 30%)	50%	Ambiguous	30	56	137
(60%, 40%)	60%	Compound	36	29	106
(70%, 30%)	70%	Compound	36	45	123
(70%, 30%)	50%	Compound	28	62	137

Table B.1: Absolute confirmation bias and absolute contradiction bias

Notes: This table shows the numbers of subjects that are classified into absolute confirmation bias and absolute contradiction bias. Only subjects who face comparable belief updating problems in the uncertain information round and its corresponding simple information round are counted.

B Additional results on the experiment

B.1 An alternative classification for behaviors in uncertain information rounds

In this section, I consider two alternative subject categories based on behaviors in uncertain information rounds: absolute confirmation bias and absolute contradiction bias. In an uncertain information round where the odds of the bets are not 50%, if a subject's uncertainty premium for the bet with higher-than-50% odds is weakly negative, her uncertainty premium for the other bet is weakly positive, and at least one of the two is not zero, then I classify this subject into the category of absolute confirmation bias. If, on the contrary, the bet with high odds has a weakly positive uncertainty premium, the other one has a weakly negative premium, and at least one is not zero, then this subject is absolute contradiction-biased. In rounds where the odds are 50-50, I do not classify subjects into these two categories.

Table B.1 shows the number of subjects in these two categories. Except in one round, there are more subjects in the category of absolute contradiction bias in every uncertain information round. This suggests that uncertainty in information accuracy does not lead to prevalent confirmation bias.

B.2 Order effects and anchor effects in the experimental data

In this section, I show that the experimental results are robust to order effects and anchor effects. Recall that in the experiment there are three different orders among parts, and two orders between

Session	Order between parts	Ambiguous block first?	Number of subjects
1	1-4-2-3-5	No	16
2	1-4-2-3-5	No	16
3	1-4-2-3-5	No	13
4	1-4-2-3-5	Yes	16
5	1-4-2-5-3	No	15
6	1-2-3-4-5	No	16
7	1-2-3-4-5	Yes	16
8	1-4-2-5-3	Yes	15
9	1-4-2-5-3	Yes	15
10	1-2-3-4-5	No	11
11	1-2-3-4-5	Yes	16

Table B.2: Description of sessions

compound and ambiguous uncertainty within parts (Table B.2). To check for order effects, I estimate the structural model separately for subjects who face each of the five orders and compare the resulting estimates for the key parameters. Table B.4 shows the result. Across different cuts of data, the α's are consistently larger than 0.5 and the ε's are consistently larger than 0. This suggests that our key results are robust to order effects.

In all three different orders among parts, Part 2 (simple prior with simple information) comes before Part 3 (simple prior with uncertain information) and Part 5 (uncertain prior with simple information). This raises the question whether subjects anchor their answers in Parts 3 and 5 to those in Part 2.

To address this concern, I first conduct a within-subject analysis by running a paired t-test between the conditional CEs in each uncertain information (prior) problem and their counterparts in the corresponding simple problem. The subjects who are included in the paired t-tests are those who receive comparable messages in the two corresponding rounds,[44] so their conditional CEs in the uncertain information (prior) round could potentially be anchored to their answers in the corresponding simple round. The other subjects who do not receive comparable messages are not subject to the anchor effect, and I compare the mean of their conditional CEs in the uncertain information (prior) problem to the mean conditional CE in the corresponding simple problem in an unpaired t-test, which is a between-subject analysis.

Table B.5 reports results for uncertain information problems. For subjects who receive com-

[44]Receiving comparable messages in two corresponding rounds means that the uncertainty premiums of the Red bet and the Blue bet in the uncertain information (prior) round can be calculated from data. See Section 4.2 and Appendix A.2 for the definition of uncertainty premiums.

	Order	Prior	Information accuracy
Part 1	1	(50%, 50%)	-
	2	(60%, 40%)	-
	3	(70%, 30%)	-
Part 2	1	(50%, 50%)	70%
	2	(60%, 40%)	60%
	3	(70%, 30%)	70%
	4	(70%, 30%)	50%
Part 3 (compound/ambiguous block)	1	(50%, 50%)	90% or 50%
	2	(60%, 40%)	90% or 30%
	3	(70%, 30%)	90% or 50%
	4	(70%, 30%)	90% or 10%
Part 4 (compound/ambiguous block)	1	(90%, 10%) or (30%, 70%)	-
	2	(90%, 10%) or (10%, 90%)	-
	3	(90%, 10%) or (50%, 50%)	-
Part 5 (compound/ambiguous block)	1	(90%, 10%) or (50%, 50%)	70%
	2	(90%, 10%) or (10%, 90%)	70%
	3	(90%, 10%) or (50%, 50%)	60%
	4	(90%, 10%) or (30%, 70%)	50%

Table B.3: Order between rounds within each part

			Ambiguity first?		Order among parts		
			No	Yes	1-2-3-4-5	1-4-2-3-5	1-4-2-5-3
α	Compound	Info accuracy	0.56 (0.03)	0.52 (0.02)	0.51 (0.02)	0.56 (0.03)	0.55 (0.03)
		Priors w/o updating	0.55 (0.02)	0.55 (0.02)	0.58 (0.02)	0.51 (0.02)	0.55 (0.02)
		Priors with updating	0.62 (0.04)	0.6 (0.05)	0.68 (0.06)	0.59 (0.05)	0.56 (0.03)
	Ambiguous	Info accuracy	0.57 (0.03)	0.6 (0.03)	0.54 (0.03)	0.62 (0.03)	0.6 (0.03)
		Priors w/o updating	0.65 (0.03)	0.56 (0.03)	0.59 (0.03)	0.61 (0.03)	0.61 (0.03)
		Priors with updating	0.63 (0.05)	0.57 (0.03)	0.59 (0.04)	0.61 (0.04)	0.59 (0.03)
ε	Compound	Info accuracy	0.22 (0.06)	0 (0.07)	0.02 (0.06)	0.05 (0.08)	0.25 (0.05)
		Priors w/o updating	0.14 (0.06)	0.17 (0.06)	0.13 (0.07)	0.25 (0.07)	0.04 (0.06)
		Priors with updating	0.06 (0.09)	0.21 (0.13)	0.13 (0.13)	0.28 (0.1)	-0.1 (0.12)
	Ambiguous	Info accuracy	0.18 (0.07)	0.17 (0.08)	0.22 (0.08)	0.02 (0.08)	0.29 (0.05)
		Priors w/o updating	0.2 (0.06)	0.32 (0.07)	0.24 (0.05)	0.27 (0.08)	0.25 (0.07)
		Priors with updating	0.15 (0.11)	-0.28 (0.14)	-0.19 (0.16)	0.05 (0.1)	-0.22 (0.1)

Table B.4: Order effects in estimates of uncertainty attitudes

Notes: This table shows the estimates of α's and ε's by the order between compound and ambiguous blocks and the order among parts. Numbers in parentheses are standard errors computed by a bootstrap clustered at the subject level.

parable messages in the uncertain information problem and the corresponding simple problem ("within-subject"), it is apparent that uncertain information leads to under-reaction to news. There is also evidence for pessimism caused by uncertain information accuracy. First, the effect sizes are more likely to be significant for good news than for bad news. Second, in half of the comparisons with neutral information, CEs conditional on uncertain information are significantly lower. (In the other comparisons with neutral information, the uncertain CEs are higher but the differences are not significant.) The results of the between-subject analysis are more noisy, but the overall patterns of under-reaction and pessimism remain present.

Table A.5 reports results for uncertain prior problems. Despite the noise in the results, in the majority of the comparisons in both within- and between-subject analysis, the conditional CEs of uncertain bets are lower than their simple counterparts, suggesting that uncertain priors in belief updating problems lead to pessimism.

Taken together, the key effects of uncertain information and uncertain priors are robust to order effects and anchor effects.

Prior and info	Type of information	within-subject $\overline{CE}(simp) - \overline{CE}(unc)$	N	between-subject $\overline{CE}(simp) - \overline{CE}(unc)$	N(simp)	N(unc)
odds=30%, accu=70%	compound	1.25 (0.11)	28	0.11 (0.93)	54	16
good news	ambiguous	2.4 (0.06)	15	-0.81 (0.45)	54	32
odds=40%, accu=60%	compound	0.8 (0.23)	59	1.19 (0.27)	91	26
good news	ambiguous	1.14 (0.12)	29	0.09 (0.93)	91	31
odds=50%, accu=70%	compound	0.52 (0.14)	163			
good news	ambiguous	1.66 (0)	164			
odds=60%, accu=60%	compound	0.55 (0.25)	47	-0.24 (0.75)	73	33
good news	ambiguous	1.9 (0)	42	3.25 (0)	73	63
odds=70%, accu=70%	compound	0.86 (0.02)	95	2.08 (0.02)	111	26
good news	ambiguous	0.54 (0.1)	79	1.74 (0.03)	111	39
odds=30%, accu=70%	compound	0.16 (0.57)	95	-0.1 (0.91)	111	26
bad news	ambiguous	-0.32 (0.4)	79	0.32 (0.67)	111	39
odds=40%, accu=60%	compound	-0.28 (0.47)	47	-1.56 (0.05)	73	33
bad news	ambiguous	-0.12 (0.83)	42	1.54 (0.02)	73	63
odds=50%, accu=70%	compound	-0.59 (0.07)	163			
bad news	ambiguous	-0.47 (0.14)	165			
odds=60%, accu=60%	compound	-0.64 (0.15)	59	-1.42 (0.15)	91	26
bad news	ambiguous	-1.1 (0.16)	29	-1.67 (0.07)	91	31
odds=70%, accu=70%	compound	-1.14 (0.22)	28	-4.05 (0)	54	16
bad news	ambiguous	-0.73 (0.56)	15	-2.7 (0.01)	54	32
odds=30%, accu=50%	compound	-0.33 (0.33)	163			
neutral news	ambiguous	-0.17 (0.66)	162			
odds=70%, accu=50%	compound	0.6 (0.05)	163			
neutral news	ambiguous	0.65 (0.03)	162			

Table B.5: Comparison between CEs of simple bets conditional on uncertain and simple information, with and without anchoring

Notes: This table shows the differences in mean conditional CEs between uncertain information problems and simple information problems. Numbers in parentheses are p-values in t-tests, and N is the number of subjects included. For example, the top row of the table states that there are 28 subjects who receive good news both in the compound information problem and in the simple information problem where the prior is 30% and (midpoint) information accuracy 70%. Among these subjects, the difference in mean conditional CEs between these two problems is \$1.25 and the p-value of the paired t-test is 0.11. Sixteen subjects receive good news in the compound information problem but not in the simple problem, and there are 54 subjects who receive good news in the simple problem in total. The difference between the mean simple conditional CE of the latter group and the mean compound conditional CE of the former is 0.11, and the p-value of the unpaired t-test is 0.93.

B.3 Inherent belief updating biases

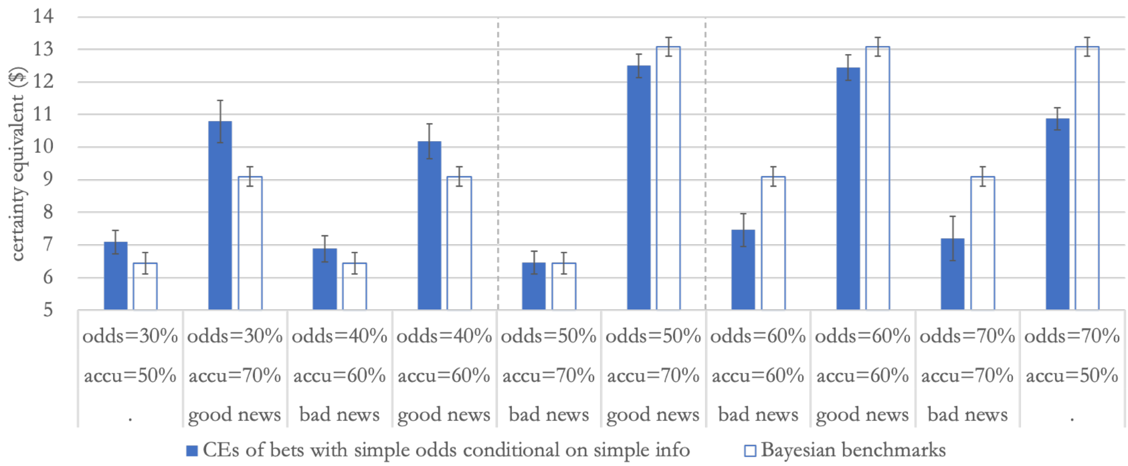

Figure B.1: Belief updating with simple priors and simple information

Notes: The figure shows the mean CEs of simple bets conditional on simple information. The horizontal axis lists the combinations of prior and information. For example, "odds=30%, accu=70%, good news" refers to simple bets with a 30% winning odds conditional on 70%-accurate simple good news. The red bars represent the mean conditional CEs and the blue bars represent the Bayesian benchmarks. The Bayesian benchmarks for "odds=30%, accu=70%, bad news" and "odds=70%, accu=70%, good news" are missing because I do not elicit CEs for simple bets whose odds match the Bayesian posteriors of these two tasks. Error bars represent +/- one standard error.

B.4 Individual-level structural estimation

In this section, I estimate the structural model in Section 6 for each individual subject. Table B.7 shows the median estimates and the percentages of subjects of whom $\alpha > 0.5$ and $\varepsilon > 0$. For all types of uncertainty, the median α is greater than 0.5 and the median ε is greater than 0. This is consistent with the pattern in the aggregate estimates, except that in the aggregate estimation, the ε for ambiguous priors (in problems with belief updating) is insignificantly negative. Table B.8 shows the median estimates of the parameters in risk preferences and inherent belief updating biases, which are rather close to the aggregate estimates in Table 6.3.

Odds	Information accuracy	Good/Bad news	Mean conditional CE	Mean Bayesian CE	Paired t-test for Bayesianism	N
30%	70%	bad	5.70			111
30%	50%		6.45	7.10	0.037	163
30%	70%	good	9.10	10.80	0.000	54
40%	60%	bad	6.45	6.89	0.244	73
40%	60%	good	9.10	10.19	0.026	90
50%	70%	bad	6.45	6.47	0.956	165
50%	70%	good	13.09	12.51	0.098	164
60%	60%	bad	9.10	7.46	0.002	90
60%	60%	good	13.09	12.45	0.034	73
70%	70%	bad	9.10	7.20	0.049	54
70%	50%		13.09	10.88	0.000	163
70%	70%	good	14.74			111

Table B.6: CEs of simple bets conditional on simple information

Notes: This table compares the mean CEs of simple bets conditional on simple information to the mean CEs given Bayesian posteriors. Numbers under paired t-tests are two-sided p-values.

Type of uncertainty		α		ε	
		Median	$\%(\alpha > 0.5)$	Median	$\%(\varepsilon > 0)$
Info accuracy	Compound	0.52	57.0%	0.23	61.8%
	Ambiguous	0.57	63.6%	0.31	69.7%
Priors (without updating)	Compound	0.54	58.8%	0.17	67.3%
	Ambiguous	0.59	69.7%	0.27	69.7%
Priors (with updating)	Compound	0.53	61.2%	0.25	63.0%
	Ambiguous	0.56	64.2%	0.19	59.4%

Table B.7: Individual estimates of α and ε

Notes: This table shows the median estimates and the percentages of subjects of whom $\alpha > 0.5$ and $\varepsilon > 0$. All estimations are conducted using nonlinear least squares.

Parameter	Median estimate
a	0.92
b	1.05
β	0.47
r_g	0.96
r_b	0.76

Table B.8: Individual estimates of incidental parameters

Notes: In the individual-level estimations, all parameters are estimated for each subject. This model is estimated using nonlinear least squares.

B.5 Alternative specifications of the structural model

B.5.1 Insensitivity proportional to the range of uncertainty

In the ε-α-maxmin model, agents put a fixed weight ε on the symmetric and maximally uncertain probability (50%, 50%) so long as there are two possible probability distributions. Alternatively, it's also possible that the degree of insensitivity depends on the range of uncertainty. In this section, I consider an alternative model where ε, the parameter for uncertainty-induced insensitivity, scales with the gap between the two possible priors or levels of information accuracy. Formally, define

$$\tilde{W}(x, y; \tilde{\varepsilon}, \alpha) := (1 - \tilde{\varepsilon} \cdot |x - y|)[(1 - \alpha)x + \alpha y] + \tilde{\varepsilon} \cdot |x - y| \cdot 0.5.$$

In the setting of my experiment where the payoff of the bet is either 1 util or 0 util, and the winning odds is either p_h or p_l with $p_h > p_l$, an agent with *scaled* ε-α-maxmin (expected utility) preference evaluates the bet by

$$u = \tilde{W}(p_h, p_l; \tilde{\varepsilon}, \alpha).$$

The parameter $\tilde{\varepsilon}$ can be interpreted as insensitivity per unit of range of uncertainty.

Like the original ε-α-maxmin preference, the scaled version can also be written in the functional form of Choquet expected utility. Hence, I can invoke Eichberger et al. (2007) to derive the conditional evaluations of bets in belief updating problems under Full Bayesian updating. Applying the same risk preference specification and generalization to Bayes' rule, I obtain the empirical model shown in Table B.9.

I estimate this model at the aggregate level using nonlinear least squares. Tables B.10 and B.11 show the estimates of the key parameters and the incidental ones, respectively. These estimates have very similar patterns as those of the original empirical model in Section 6.

Part	Prior	Information	Model for CE
1	Simple	No information	$CE(p) = M^{Prelec}(p)$
2	Simple	Simple	$CE(p, m, \psi) = M^{Prelec}\left(Pr^{GB}(G\|p, m, \psi)\right)$
3	Simple	Uncertain	$CE(p, g, \psi_h \text{ or } \psi_l) = M^{Prelec}\left(Pr^{GB}(G\|p, g, \tilde{W}(\psi_h, \psi_l; \tilde{\varepsilon}, \alpha))\right)$ $CE(p, b, \psi_h \text{ or } \psi_l) = M^{Prelec}\left(Pr^{GB}(G\|p, b, \tilde{W}(\psi_l, \psi_h; \tilde{\varepsilon}, \alpha))\right)$
4	Uncertain	No information	$CE(p_h \text{ or } p_l) = M^{Prelec}\left(\tilde{W}(p_h, p_l; \tilde{\varepsilon}, \alpha)\right)$
5	Uncertain	Simple	$CE(p_h \text{ or } p_l, m, \psi) = M^{Prelec}\left(Pr^{GB}(G\|\tilde{W}(p_h, p_l; \tilde{\varepsilon}, \alpha), m, \psi)\right)$

Table B.9: Summary of the empirical model based on the scaled ε-α-maxmin preference

Type of uncertainty		α	$\tilde{\varepsilon}$
Info accuracy	Compound	0.54 (0.02)	0.20 (0.11)
	Ambiguous	0.59 (0.02)	0.38 (0.12)
Priors (without updating)	Compound	0.55 (0.02)	0.31 (0.09)
	Ambiguous	0.63 (0.03)	0.55 (0.09)
Priors (with updating)	Compound	0.62 (0.04)	0.33 (0.19)
	Ambiguous	0.60 (0.03)	-0.15 (0.23)

Table B.10: Aggregate-level estimates of α and $\tilde{\varepsilon}$ in the model based on the scaled ε-α-maxmin preference

Notes: All parameters are assumed to be homogeneous among subjects. Numbers in the parentheses are standard errors, which are computed by a bootstrap clustered at the subject level. The model is estimated using nonlinear least squares.

Parameter	Estimate (s.e.)
a	0.77 (0.04)
b	1.03 (0.03)
β	0.52 (0.04)
r_g	0.97 (0.06)
r_b	0.72 (0.06)

Table B.11: Aggregate-level estimates of incidental parameters in the model based on the scaled ε-α-maxmin preference

Notes: All parameters are assumed to be homogeneous among subjects. Numbers in parentheses are standard errors, which are computed by a bootstrap clustered at the subject level. The model is estimated using nonlinear least squares.

Part	Prior	Information	Model for CE
1	Simple	No information	$CE(p) = M^{Prelec}(p)$
2	Simple	Simple	$CE(p, m, \psi) = M^{Prelec}\left(Pr^{GB}(G\|p, m, \psi)\right)$
3	Simple	Uncertain	$CE(p, g, \psi_h \text{ or } \psi_l) = M^{Prelec}\left(Pr^{GB}(G\|p, g, \hat{W}(\psi_h, \psi_l; \varepsilon, \alpha, \delta))\right)$ $CE(p, b, \psi_h \text{ or } \psi_l) = M^{Prelec}\left(Pr^{GB}(G\|p, b, \hat{W}(\psi_l, \psi_h; \varepsilon, \alpha, \delta))\right)$
4	Uncertain	No information	$CE(p_h \text{ or } p_l) = M^{Prelec}\left(\hat{W}(p_h, p_l; \varepsilon, \alpha, \delta)\right)$
5	Uncertain	Simple	$CE(p_h \text{ or } p_l, m, \psi) = M^{Prelec}\left(Pr^{GB}(G\|\hat{W}(p_h, p_l; \varepsilon, \alpha, \delta), m, \psi)\right)$

Table B.12: Summary of the empirical model based on the ε-α-δ-maxmin preference

B.5.2 Alternative baseline probabilities

In this section, I generalize the original ε-α-maxmin model to allow the baseline probability that receives ε weight to be a free parameter rather than (50%, 50%). Formally, define

$$\hat{W}(x, y; \varepsilon, \alpha, \delta) := (1 - \varepsilon)[(1 - \alpha)x + \alpha y] + \varepsilon \cdot \delta.$$

In the setting of my experiment where the payoff of the bet is either 1 util or 0 util, and the winning odds is either p_h or p_l with $p_h > p_l$, an agent with *ε-α-δ-maxmin (expected utility) preference* evaluates the bet by

$$u = \hat{W}(p_h, p_l; \varepsilon, \alpha, \delta).$$

This model can be written in the functional form of Choquet expected utility, so I can apply Eichberger et al. (2007) to obtain its Full Bayesian updating extension to problems with belief updating. With the same risk preference specification and generalization to Bayes' rule, Table B.12 shows the empirical model based on the ε-α-δ-maxmin preference.

I estimate this model using nonlinear least squares at the aggregate level. The estimates of ε, α and δ are shown in Table B.13. The point estimates of δ are not significantly different from 0.5 except for compound priors in problems without updating. Moreover, the estimates of α and ε have similar patterns as those under the original empirical model. This suggests that the main results of the structural estimation is robust to allowing the baseline probability which receives ε weight to be a free parameter.

Type of uncertainty		α	ε	δ
Info accuracy	Compound	0.55 (0.04)	0.11 (0.05)	0.59 (2.55)
	Ambiguous	0.6 (0.04)	0.18 (0.05)	0.55 (0.21)
Priors (without updating)	Compound	0.61 (0.04)	0.15 (0.04)	0.72 (0.17)
	Ambiguous	0.63 (0.05)	0.25 (0.05)	0.55 (0.08)
Priors (with updating)	Compound	0.63 (0.06)	0.14 (0.09)	0.61 (7.45)
	Ambiguous	0.6 (0.05)	-0.08 (0.09)	0.42 (35.1)

Table B.13: Aggregate estimates of α, ε and δ in the model based on the ε-α-δ-maxmin preference

Notes: All parameters are assumed to be homogeneous among subjects. Numbers in the parentheses are standard errors, which are computed by a bootstrap clustered at the subject level. The model is estimated using nonlinear least squares.

C Additional results on the ε-α-maxmin EU preferences

C.1 Representing ε-α-maxmin preferences using Choquet integrals

Under Choquet expected utility (CEU) (Schmeidler, 1989), each event $E \subseteq S$ is assigned a value called *capacity* by a set function $v : 2^S \to [0, 1]$. A capacity satisfies the following two conditions:

- $v(\emptyset) = 0$ and $v(S) = 1$;
- If $E_1 \subseteq E_2$, then $v(E_1) \leq v(E_2)$.

Capacities generalize probability measures by allowing measures of sets to be non-additive. Given a capacity and a vNM utility function for simple lotteries, a CEU agent evaluates an act by taking the Choquet integral of utility. Formally, denote by $u(s)$ the vNM utility of the lottery assigned to state $s \in S$, then the CEU evaluation of the act is

$$\int_{\mathbb{R}} v\left(\{s \in S | u(s) \geq t\}\right) dt.$$

Now I show the CEU representation of the ε-α-maxmin preferences. I construct the capacity of each event so that it equals the ε-α-maxmin evaluation of a bet that pays out 1 util if this event happens and 0 otherwise. Recall that $W(x, y; \varepsilon, \alpha) = (1 - \varepsilon)[(1 - \alpha)x + \alpha y] + \varepsilon \cdot 0.5$. In a problem with uncertain prior (p_h or p_l), $p_h > p_l$, and no information, let the state space be $\{G, B\}$. Then the capacity of each event is

- $v(\{G\}) = W(p_h, p_l; \varepsilon, \alpha)$,

- $\nu(\{B\}) = W(1 - p_l, 1 - p_h; \varepsilon, \alpha)$.

In a problem with simple prior p and uncertain information accuracy (ψ_h or ψ_l), $\psi_h > \psi_l$, let the state space be $\{Gg, Gb, Bg, Bb\}$, where the capital letter represents the true outcome of the bet and the lower-case letter the realized message.[45] The capacity of each event is

- $\nu(\{Gg\}) = p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha)$,
- $\nu(\{Gb\}) = p \cdot W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha)$,
- $\nu(\{Bg\}) = (1 - p) \cdot W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha)$,
- $\nu(\{Bb\}) = (1 - p) \cdot W(\psi_h, \psi_l; \varepsilon, \alpha)$,
- $\nu(\{Gg, Gb\}) = p$,
- $\nu(\{Bg, Bb\}) = 1 - p$,
- $\nu(\{Gg, Bb\}) = W(\psi_h, \psi_l; \varepsilon, \alpha)$,
- $\nu(\{Gb, Bg\}) = W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha)$,
- $\nu(\{Gg, Bg\}) = \begin{cases} p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha) + (1 - p) \cdot W(1 - \psi_h, 1 - \psi_l; \varepsilon, \alpha), & \text{if } p \geq 0.5 \\ p \cdot W(\psi_l, \psi_h; \varepsilon, \alpha) + (1 - p) \cdot W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha), & \text{if } p < 0.5 \end{cases}$,
- $\nu(\{Gb, Bb\}) = \begin{cases} p \cdot W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha) + (1 - p) \cdot W(\psi_l, \psi_h; \varepsilon, \alpha), & \text{if } p \geq 0.5 \\ p \cdot W(1 - \psi_h, 1 - \psi_l; \varepsilon, \alpha) + (1 - p) \cdot W(\psi_h, \psi_l; \varepsilon, \alpha), & \text{if } p < 0.5 \end{cases}$,
- $\nu(\{Gg, Gb, Bg\}) = p + (1 - p) \cdot W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha)$,
- $\nu(\{Gg, Gb, Bb\}) = p + (1 - p) \cdot W(\psi_h, \psi_l; \varepsilon, \alpha)$,
- $\nu(\{Gg, Bg, Bb\}) = p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha) + 1 - p$,
- $\nu(\{Gb, Bg, Bb\}) = p \cdot W(1 - \psi_l, 1 - \psi_h; \varepsilon, \alpha) + 1 - p$.

In a problem with uncertain prior (p_h or p_l), $p_h > p_l$, and simple information accuracy ψ, let the state space be $\{Gg, Gb, Bg, Bb\}$, where the capital letter represents the true outcome of the bet and the lower-case letter the realized message. The capacity of each event is

[45] The definition of state space is natural because it is the coarsest common refinement of the partition generated by the events with uncertain probabilities $\{\{Gg, Bb\}, \{Gb, Bg\}\}$ and the partition generated by the payoffs $\{\{Gg, Gb\}, \{Bg, Bb\}\}$.

- $\nu(\{Gg\}) = W(p_h, p_l; \varepsilon, \alpha) \cdot \psi$,
- $\nu(\{Gb\}) = W(p_h, p_l; \varepsilon, \alpha) \cdot (1 - \psi)$,
- $\nu(\{Bg\}) = W(1 - p_l, 1 - p_h; \varepsilon, \alpha) \cdot (1 - \psi)$,
- $\nu(\{Bb\}) = W(1 - p_l, 1 - p_h; \varepsilon, \alpha) \cdot \psi$,
- $\nu(\{Gg, Gb\}) = W(p_h, p_l; \varepsilon, \alpha)$,
- $\nu(\{Bg, Bb\}) = W(1 - p_l, 1 - p_h; \varepsilon, \alpha)$,
- $\nu(\{Gg, Bb\}) = \psi$,
- $\nu(\{Gb, Bg\}) = 1 - \psi$,
- $\nu(\{Gg, Bg\}) = \begin{cases} W(p_h, p_l; \varepsilon, \alpha) \cdot \psi + W(1 - p_h, 1 - p_l; \varepsilon, \alpha) \cdot (1 - \psi), & \text{if } p_h + p_l \geq 1 \\ W(p_l, p_h; \varepsilon, \alpha) \cdot \psi + W(1 - p_l, 1 - p_h; \varepsilon, \alpha) \cdot (1 - \psi), & \text{if } p_h + p_l < 1 \end{cases}$,
- $\nu(\{Gb, Bb\}) = \begin{cases} W(p_l, p_h; \varepsilon, \alpha) \cdot (1 - \psi) + W(1 - p_l, 1 - p_h; \varepsilon, \alpha) \cdot \psi, & \text{if } p_h + p_l \geq 1 \\ W(p_h, p_l; \varepsilon, \alpha) \cdot (1 - \psi) + W(1 - p_h, 1 - p_l; \varepsilon, \alpha) \cdot \psi, & \text{if } p_h + p_l < 1 \end{cases}$,
- $\nu(\{Gg, Gb, Bg\}) = W(p_h, p_l; \varepsilon, \alpha) \cdot \psi + 1 - \psi$,
- $\nu(\{Gg, Gb, Bb\}) = W(p_h, p_l; \varepsilon, \alpha) \cdot (1 - \psi) + \psi$,
- $\nu(\{Gg, Bg, Bb\}) = W(1 - p_l, 1 - p_h; \varepsilon, \alpha) \cdot (1 - \psi) + \psi$,
- $\nu(\{Gb, Bg, Bb\}) = W(1 - p_l, 1 - p_h; \varepsilon, \alpha) \cdot \psi + 1 - \psi$.

C.2 Proofs of results on ε-α-maxmin preferences

Proof of Propositions 1 and 5. Eichberger et al. (2007) defines Full Bayesian updating for capacities as follows. The capacity of event A conditional on realized message E is

$$\nu(A|E) = \frac{\nu(A \cap E)}{\nu(A \cap E) + 1 - \nu(A \cup E^c)}.$$

We can obtain the Full Bayesian conditional evaluations of bets by directly applying the definition above to the CEU representation of ε-α-maxmin preferences. For example, in an uncertain information problem, the capacity of G conditional on message g is

$$\nu(\{Gg\}|\{Gg, Bg\}) = \frac{\nu(\{Gg\})}{\nu(\{Gg\}) + 1 - \nu(\{Gg, Gb, Bb\})}$$

$$= \frac{p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha)}{p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha) + (1-p) \cdot (1 - W(\psi_h, \psi_l; \varepsilon, \alpha))}$$
$$= Pr^{Bayes}(G|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha)).$$

Hence, the evaluation of the bet conditional on message g is

$$u = (1-\nu(\{Gg\}|\{Gg, Bg\}))\cdot 0+\nu(\{Gg\}|\{Gg, Bg\})\cdot 1 = \nu(\{Gg\}|\{Gg, Bg\}) = Pr^{Bayes}(G|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha)).$$

The conditional evaluation given message b and those in uncertain prior problems can be similarly derived.

The comparative statics of the conditional evaluations with respect to α and ε are straightforward.

Proof of Propositions 2 and 6. Under Dynamically consistent updating (Hanany and Klibanoff, 2007), the agent forms a contingent plan of actions before the message realizes and executes the plan resolutely after observing the message. In our example where the agent chooses between a bet and a sure amount of utils, the contingent plan, $a = (a(g), a(b))$, specifies an action $a(m) \in \{Bet, Sure\}$ conditional on good news and bad news. Let $U(a(m), E)$ denote the utility of action $a(m)$ under payoff-relevant event E. The optimal plan maximizes utility from the ex-ante perspective. In an uncertain information problem, the ex-ante utility of an agent with an ε-α-maxmin preference is

$$W(\psi_h, \psi_l; \varepsilon, \alpha)\cdot[p\cdot U(a(g), G)+(1-p)\cdot U(a(b), B)]+(1-W(\psi_h, \psi_l; \varepsilon, \alpha))\cdot[p\cdot U(a(b), G)+(1-p)\cdot U(a(g), B)]$$

if her plan of action is (Bet, Bet), $(Bet, Sure)$ or $(Sure, Sure)$ and

$$W(\psi_l, \psi_h; \varepsilon, \alpha)\cdot[p\cdot U(a(g), G)+(1-p)\cdot U(a(b), B)]+(1-W(\psi_l, \psi_h; \varepsilon, \alpha))\cdot[p\cdot U(a(b), G)+(1-p)\cdot U(a(g), B)]$$

if her plan is $(Sure, Bet)$.

It's straightforward that the payoff of (Bet, Bet) equals p and that of $(Sure, Sure)$ equals the sure amount u. Note that both payoffs are independent from the information accuracy. The intuition is that if the agent's action is unaffected by the realization of information, then the ex-ante utility is not exposed to the uncertainty in the information. This is in contrast with the ex-post utility conditional on the realized message. The only choice that makes the ex-post conditional evaluation independent from the uncertainty in the information is choosing the sure amount of utils.

Since $W(\psi_h, \psi_l; \varepsilon, \alpha) \geq 1 - W(\psi_l, \psi_h; \varepsilon, \alpha)$, it can be shown by simple algebra that $(Sure, Bet)$ always leads to lower ex-ante utility than $(Bet, Sure)$. Hence, I only need to consider (Bet, Bet), $(Sure, Sure)$ and $(Bet, Sure)$ as the candidate optimal plans.

We know that $(Sure, Sure)$ yields a higher utility than (Bet, Bet) if and only if $u > p$. Hence, to pin down the optimal plan for each u, we only need to find the u such that $(Bet, Sure)$ is optimal. The plan $(Bet, Sure)$ yields a higher utility than $(Sure, Sure)$ if and only if

$$W(\psi_h, \psi_l; \varepsilon, \alpha) \cdot p \cdot (1-u) - (1 - W(\psi_h, \psi_l; \varepsilon, \alpha))(1-p) \cdot u > 0$$

$$\iff u < Pr^{Bayes}(G|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha)).$$

Similarly, $(Bet, Sure)$ yields a higher utility than (Bet, Bet) if and only if

$$u > Pr^{Bayes}(G|p, b, W(\psi_h, \psi_l; \varepsilon, \alpha)).$$

The two inequalities can be simultaneously satisfied if and only if $(1-\alpha)\psi_h + \alpha\psi_l > 0.5$. When this condition holds, it's easy to check that $(Bet, Sure)$ is indeed optimal in the interval between the two right-hand side expressions. If we interpret the upper and lower boundaries of the interval in which $(Bet, Sure)$ is optimal as the "conditional evaluations" given good news and bad news, respectively, then these "conditional evaluations" coincide exactly with the Bayesian conditional evaluations with the information accuracy being $W(\psi_h, \psi_l; \varepsilon, \alpha)$.

If $(1-\alpha)\psi_h + \alpha\psi_l < 0.5$, then there is no u such that $(Bet, Sure)$ is optimal. Hence, the agent's optimal plan is to not respond to the information at all: she always chooses the sure amount of utils if $u > p$ and always takes the bet if $u < p$, *regardless of the realized message*.

In an uncertain prior problem, the ex-ante utility of an agent with an ε-α-maxmin preference is

$$W(p_h, p_l; \varepsilon, \alpha)\cdot[\psi\cdot U(a(g), G)+(1-\psi)\cdot U(a(b), G)]+(1-W(p_h, p_l; \varepsilon, \alpha))\cdot[(1-\psi)\cdot U(a(g), B)+(1-\psi)\cdot U(a(b), B)].$$

The ex-ante expected utility of $(Sure, Sure)$ is still u but that of (Bet, Bet) is now $W(p_h, p_l; \varepsilon, \alpha)$. The ex-ante expected utility of $(Bet, Sure)$ is again always higher than that of $(Sure, Bet)$.

We know that (Bet, Bet) yields a higher ex-ante payoff than $(Sure, Sure)$ if $u > W(p_h, p_l; \varepsilon, \alpha)$. Simple algebra shows that $(Bet, Sure)$ yields a higher payoff than $(Sure, Sure)$ if and only if

$$u < Pr^{Bayes}(G|W(p_h, p_l; \varepsilon, \alpha), g, \psi)$$

and $(Bet, Sure)$ yields a higher payoff than (Bet, Bet) if and only if

$$u > Pr^{Bayes}(G|W(p_l, p_h; \varepsilon, \alpha), b, \psi).$$

The two inequalities above can always be compatible. Note that the two expressions on the right-hand side are exactly the same as the conditional evaluations in the same uncertain prior problem under Full Bayesian updating. This suggests that for uncertain prior problems, Full Bayesian updating and Dynamically consistent updating make the same predictions under ε-α-maxmin preferences.

The comparative statics of conditional evaluations with respect to ε and α are straightforward.

Proof of Propositions 3 and 7. In an uncertain information problem, the likelihood of message g is $p \cdot \psi + (1-p) \cdot (1-\psi)$. If $p > 0.5$, then the likelihood is increasing in ψ and thus ψ_h is selected. If $p < 0.5$, then ψ_l is selected upon the realization of g. Similarly, the likelihood of message b is $p \cdot (1-\psi) + (1-p) \cdot \psi$. If $p > 0.5$, then ψ_l is selected and if $p < 0.5$, ψ_h is selected. If $p = 0.5$, then both ψ_h and ψ_l are retained regardless of the realized message. Uncertain prior problems are analogous.

C.3 Updating by proxy

Gul and Pesendorfer (2018) introduces an updating rule based on the idea that the realization of information provides ex-post randomization which hedges against uncertainty (in priors or information accuracy) for pessimistic agents. Their rule, termed *updating by proxy*, only applies to Choquet EU preferences with totally monotone capacity. A capacity ν is totally monotone if for all events E_1 and E_2, $\nu(E_1) + \nu(E_2) \leq \nu(E_1 \cup E_2) + \nu(E_1 \cap E_2)$. The capacity that represents an ε-α-maxmin preference in an uncertain information problem is generically not totally monotone. For example, in an uncertain information problem, so long as $\alpha > 0.5$, we have

$$\begin{aligned} \nu(\{Gg\}) + \nu(\{Gg, Gb, Bg\}) &= p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha) + p + (1-p) \cdot W(1-\psi_l, 1-\psi_h; \varepsilon, \alpha) \\ &< p \cdot W(\psi_l, \psi_h; \varepsilon, \alpha) + p + (1-p) \cdot W(1-\psi_l, 1-\psi_h; \varepsilon, \alpha) \\ &= \nu(\{Gg, Bg\}) + \nu(\{Gg, Gb\}), \end{aligned}$$

which violates total monotonicity.

However, a modified version of my setting can be represented by a totally monotone capacity. For example, in an uncertain information problem, define capacity $\tilde{\nu}$ such that

$$\tilde{\nu}(\{Gg, Bg\}) = p \cdot W(\psi_h, \psi_l; \varepsilon, \alpha) + (1-p) \cdot W(1-\psi_l, 1-\psi_h; \varepsilon, \alpha),$$

$$\tilde{\nu}(\{Gb, Bb\}) = p \cdot W(1-\psi_l, 1-\psi_h; \varepsilon, \alpha) + (1-p) \cdot W(\psi_h, \psi_l; \varepsilon, \alpha),$$

and $\tilde{\nu}(E) = \nu(E)$ for all the other events. It's straightforward to check that $\tilde{\nu}$ is a totally monotone capacity. The intuition is that $\tilde{\nu}$ represents a setting where the information accuracy levels conditional on G and B can be different and each of them can be either ψ_h or ψ_l.

Under updating by proxy, the capacity of event E conditional on event F is

$$\nu^{UbP}(E|F) := \frac{\rho_\nu^{E\cup F^c}(E\cap F)}{\rho_\nu^S(F)},$$

where $\rho_\nu^C(A) = \sum_{s\in A}\rho_\nu^C(s)$ and $\rho_\nu^C(s)$ is the Shapley value of s in the "cooperative game" C. For details, see Gul and Pesendorfer (2018). After a few steps of algebra, it can be derived that a (weakly) pessimistic agent ($\alpha \geq 0.5$) who uses updating by proxy evaluates the bet by

$$u = \nu^{UbP}(\{Gm\}|\{Gm, Bm\}) = Pr^{Bayes}(G|p, m, W(\psi_h, \psi_l; \varepsilon, 0.5))$$

conditional on realized message $m \in \{g, b\}$.

Intuitively, a pessimistic agent no longer weights different levels of accuracy differently because the uncertainty is hedged against after the realization of information. The only remaining deviation from simple information is the ε weight on 0.5 caused by insensitivity.

In a similarly modified uncertain prior problem, a (weakly) pessimistic agent who uses updating by proxy evaluates the bet by

$$u = Pr^{Bayes}(G|W(p_h, p_l; \varepsilon, 0.5), m, \psi)$$

conditional on realized message $m \in \{g, b\}$. The following proposition summarizes the results.

Proposition 8 *Suppose a (weakly) pessimistic ε-α-maxmin agent ($\alpha \geq 0.5$) uses updating by proxy. In a modified uncertain information problem,*

1. *if $\varepsilon = 0$ and $\alpha = 0.5$, then her conditional evaluations coincide with the Bayesian evaluations conditional on information with accuracy level $\frac{\psi_h+\psi_l}{2}$;*
2. *as ε, the measure of insensitivity, increases, the conditional evaluations become closer to p;*
3. *the measure of pessimism α does not affect conditional evaluations.*

In a modified uncertain prior problem,

1. *if $\varepsilon = 0$ and $\alpha = 0.5$, then her conditional evaluations coincide with the Bayesian evaluations given simple prior $\frac{p_h+p_l}{2}$;*
2. *as ε, the measure of insensitivity, increases, the conditional evaluations become closer to ψ given good news and $1-\psi$ given bad news;*
3. *the measure of pessimism α does not affect conditional evaluations.*

D An axiomatic foundation of ε-α-maxmin EU

In this appendix, I provide an axiomatic foundation for the ε-α-maxmin expected utility preference. I only axiomatize the model in a restricted domain that closely resembles my experiment. A full-fledged revealed preference analysis of the model in a more general setting is left for future research.

There are two events, E and E^c. An act assigns each event a simple lottery. Agents' evaluations of simple lotteries satisfy the von Neumann-Morgenstern (vNM) expected utility axioms, so I identify each simple lottery with its expected utility. Assume that utility is bounded and normalize the range to the interval $[0, 1]$. An act also assigns one or two probability distributions to the events. So in this sense, the act is objective. A simple act is denoted as $(p; v_1, v_2)$, where p is the (single) probability of E, and v_1 and v_2 are the utility of the two simple lotteries assigned to E and E^c, respectively. An uncertain act is denoted as $(p_1 \text{ or } p_2; v_1, v_2)$, where p_1 and p_2 are the two possible probabilities of E. I impose no order on p_1 and p_2; in other words, "p_1 or p_2" is the same as "p_2 or p_1." I also allow for the possibility that $p_1 = p_2$. v_1 and v_2 have the same meanings as in simple acts.

Here are a couple of remarks on the framework. The framework is adapted from the objective ambiguity framework of Olszewski (2007). There are two differences. First, I treat the simple act $(p; v_1, v_2)$ and the uncertain act $(p \text{ or } p; v_1, v_2)$ as different entities. This is intended to capture the idea that multiple probabilities can cause confusion or inattention, even when the multiple probabilities are the same. Second, I make explicit the two events E and E^c and only allow for multiple probabilities on these two events (but not within the lottery assigned to each event). This is to stay close to my experiment where there are only uncertain probabilities from one source of uncertainty.

Now I introduce the axioms.

Axiom 1 (Simple vNM)

$$(p; v_1, v_2) \sim (1; pv_1 + (1-p)v_2, 0).$$

Axiom 2 (Event C-Independence) *If* $(p_1 \text{ or } p_2; v_1, v_2) \succsim (q_1 \text{ or } q_2; v_1, v_2)$*, then for all* $k \in [0, 1]$ *and* $l \in [0, 1]$

$$(kp_1 + (1-k)l \text{ or } kp_2 + (1-k)l; v_1, v_2) \succsim (kq_1 + (1-k)l \text{ or } kq_2 + (1-k)l; v_1, v_2).$$

Axiom 3 (Prize C-Independence) *If* $(p_1 \text{ or } p_2; v_1, v_2) \succsim (q_1 \text{ or } q_2; w_1, w_2)$*, then for all* $k \in [0, 1]$ *and* $l \in [0, 1]$

$$(p_1 \text{ or } p_2; kv_1 + (1-k)l, kv_2 + (1-k)l) \succsim (q_1 \text{ or } q_2; kw_1 + (1-k)l, kw_2 + (1-k)l).$$

Axiom 4 (Simple-Uncertain Independence) *If* $(1; v_1, 0) \sim (1 \text{ or } 0; w_1, 0)$ *and* $(1; v_2, 0) \sim (1 \text{ or } 0; w_2, 0)$*, then* $(1; kv_1 + (1-k)v_2, 0) \sim (1 \text{ or } 0; kw_1 + (1-k)w_2, 0)$ *for all* $k \in [0, 1]$.

Axiom 5 (Event Continuity) *For all* p_1 *and* p_2*,* $\{(q_1, q_2) | (q_1 \text{ or } q_2; v_1, v_2) \succ (p_1 \text{ or } p_2; v_1, v_2)\}$ *and* $\{(q_1, q_2) | (q_1 \text{ or } q_2; v_1, v_2) \prec (p_1 \text{ or } p_2; v_1, v_2)\}$ *are open in* $[0, 1]^2$.

Axiom 6 (Prize Continuity) *For all* v_1 *and* v_2*,* $\{(w_1, w_2) | (1 \text{ or } 0; w_1, w_2) \succ (1 \text{ or } 0; v_1, v_2)\}$ *and* $\{(w_1, w_2) | (1 \text{ or } 0; w_1, w_2) \prec (1 \text{ or } 0; v_1, v_2)\}$ *are open in* $[0, 1]^2$.

Axiom 7 (Event Monotonicity) *If* $v_1 > v_2$ *and* $p > q$*, then* $(p; v_1, v_2) > (q; v_1, v_2)$ *and* $(p \text{ or } l; v_1, v_2) \geq (q \text{ or } l; v_1, v_2)$ *for all* $l \in [0, 1]$.

Axiom 8 (Prize Monotonicity) *If* $v > w$*, then* $(1; v, l) \succ (1; w, l)$ *and* $(1 \text{ or } 0; v, l) \succsim (1 \text{ or } 0; w, l)$ *for all* $l \in [0, 1]$.

Axiom 9 (Symmetric Events)

$$(p_1 \text{ or } p_2; v_1, v_2) \sim (1 - p_1 \text{ or } 1 - p_2; v_2, v_1)$$

for all $p_1, p_2 \in [0, 1]$ *and* $v_1, v_2 \in [0, 1]$.

Axiom 10 (Constant Act Equivalence)

$$(p; v, v) \sim (p_1 \text{ or } p_2; v, v)$$

for all $p, p_1, p_2 \in [0, 1]$ *and* $v \in [0, 1]$.

Axiom 11 (Insensitivity) $(0; 1, 0) \precsim (0 \text{ or } 0; 1, 0)$ *and* $(1; 1, 0) \succsim (1 \text{ or } 1; 1, 0)$.

Axiom 12 (Centering) $(0.5; v_1, v_2) \sim (0.5 \text{ or } 0.5; v_1, v_2)$.

Theorem 1 *A preference relation on simple and uncertain acts* $\succsim$ *satisfies Axioms 1-12 if and only if there exist* $\alpha \in [0, 1]$ *and* $\varepsilon \in [0, 1]$ *such that the preference can be represented by the following utility function:*

$$U(p; v_1, v_2) = pv_1 + (1 - p)v_2,$$

$$U(p_1 \text{ or } p_2; v_1, v_2) = [(1-\varepsilon)[(1-\alpha)p_1 + \alpha p_2] + \varepsilon \cdot 0.5]v_1 + (1 - [(1-\varepsilon)[(1-\alpha)p_1 + \alpha p_2] + \varepsilon \cdot 0.5])v_2,$$

assuming without loss of generality that $p_1 v_1 + (1 - p_1)v_2 \geq p_2 v_1 + (1 - p_2)v_2$.

Proof By Prize Monotonicity and Simple vNM, the preference over simple acts can be represented as stated in the Theorem.

By Insensitivity, Centering and Event Monotonicity,

$$(0; 1, 0) \precsim (0 \text{ or } 0; 1, 0) \precsim (0.5 \text{ or } 0.5; 1, 0) \sim (0.5; 1, 0).$$

By Event Continuity and Event Monotonicity, there exists $\varepsilon \in [0, 1]$ such that

$$(0 \text{ or } 0; 1, 0) \sim (\varepsilon \cdot 0.5; 1, 0).$$

Similarly, there exists $\varepsilon' \in [0, 1]$ such that

$$(1 \text{ or } 1; 1, 0) \sim (1 - \varepsilon' \cdot 0.5; 1, 0).$$

By Simple-Uncertain Independence,

$$(0.5 \text{ or } 0.5; 1, 0) \sim (0.5 + 0.5 \cdot \frac{\varepsilon - \varepsilon'}{2}; 1, 0).$$

But since $(0.5 \text{ or } 0.5; 1, 0) \sim (0.5; 1, 0)$ by Centering, we have $\varepsilon' = \varepsilon$ by Event Monotonicity. Hence by Simple-Uncertain Independence,

$$(p \text{ or } p; 1, 0) \sim ((1 - \varepsilon)p + \varepsilon \cdot 0.5; 1, 0). \tag{2}$$

Now consider the uncertain act $(1 \text{ or } 0; 1, 0)$. By Event Monotonicity,

$$(1 \text{ or } 1; 1, 0) \succsim (1 \text{ or } 0; 1, 0) \succsim (0 \text{ or } 0; 1, 0).$$

By Event Continuity and Event Monotonicity, there exists $\alpha \in [0, 1]$ such that

$$(1 \text{ or } 0; 1, 0) \sim (1 - \alpha \text{ or } 1 - \alpha; 1, 0).$$

Now let us consider the uncertain act $(p_1 \text{ or } p_2; 1, 0)$, assuming $p_1 \geq p_2$. Since $(1 \text{ or } 0; 1, 0) \sim (1 - \alpha \text{ or } 1 - \alpha; 1, 0)$, by Event C-Independence,

$$(k \text{ or } 0; 1, 0) \sim (k(1 - \alpha) \text{ or } k(1 - \alpha); 1, 0), \tag{3}$$

$$(1 \text{ or } k; 1, 0) \sim (1 - \alpha + k\alpha \text{ or } 1 - \alpha + k\alpha; 1, 0). \tag{4}$$

If $(1-p_1)/p_2 \geq \alpha/(1-\alpha)$, then substitute $k = \frac{\alpha}{1-\alpha}p_2 + p_1$ into (3) and we get

$$(\frac{\alpha}{1-\alpha}p_2 + p_1 \text{ or } 0; 1, 0) \sim (\alpha p_2 + (1-\alpha)p_1 \text{ or } \alpha p_2 + (1-\alpha)p_1; 1, 0).$$

It's easy to show that the point (p_1, p_2) is a convex combination of $(\frac{\alpha}{1-\alpha}p_2 + p_1, 0)$ and $(\alpha p_2 + (1-\alpha)p_1, \alpha p_2 + (1-\alpha)p_1)$. Hence, again by Event C-Independence,

$$(p_1 \text{ or } p_2; 1, 0) \sim (\alpha p_2 + (1-\alpha)p_1 \text{ or } \alpha p_2 + (1-\alpha)p_1; 1, 0).$$

If $(1-p_1)/p_2 < \alpha/(1-\alpha)$, then substitute $k = p_2 - \frac{1-\alpha}{\alpha}(1-p_1)$ into (4) and we get

$$(1 \text{ or } p_2 - \frac{1-\alpha}{\alpha}(1-p_1); 1, 0) \sim (p_2 + (1-\alpha)p_1 \text{ or } p_2 + (1-\alpha)p_1; 1, 0).$$

Since the point (p_1, p_2) is a convex combination of $(\frac{\alpha}{1-\alpha}p_2 + p_1, 0)$ and $(\alpha p_2 + (1-\alpha)p_1, \alpha p_2 + (1-\alpha)p_1)$, by Event C-Independence,

$$(p_1 \text{ or } p_2; 1, 0) \sim (\alpha p_2 + (1-\alpha)p_1 \text{ or } \alpha p_2 + (1-\alpha)p_1; 1, 0). \tag{5}$$

The same steps can be analogously applied to the uncertain acts $(p_1 \text{ or } p_2; 0, 0)$, $(p_1 \text{ or } p_2; 1, 1)$ and $(p_1 \text{ or } p_2; 1, 0)$ to determine the v such that $(p_1 \text{ or } p_2; v_1, v_2) \sim (p_1 \text{ or } p_2; v, v)$ for any $0 \leq p_2 \leq p_1 \leq 1$ and $0 \leq v_2 \leq v_1 \leq 1$. By Prize Monotonicity and Prize Continuity, there exists $\beta \in [0, 1]$ such that

$$(p_1 \text{ or } p_2; 1, 0) \sim (p_1 \text{ or } p_2; 1-\beta, 1-\beta).$$

Applying Prize C-Independence and Prize Monotonicity in a similar way as in the previous paragraph, we obtain

$$(p_1 \text{ or } p_2; v_1, v_2) \sim (p_1 \text{ or } p_2; \beta v_2 + (1-\beta)v_1, \beta v_2 + (1-\beta)v_1).$$

Now I solve for β. By the indifference relations (5) and (2),

$$(p_1 \text{ or } p_2; 1, 0) \sim ((1-\varepsilon)[\alpha p_2 + (1-\alpha)p_1] + \varepsilon \cdot 0.5; 1, 0).$$

By Constant Act Equivalence and Simple vNM,

$$(p_1 \text{ or } p_2; \beta, \beta) \sim (1; \beta, \beta) \sim (\beta; 1, 0).$$

Since $(p_1 \text{ or } p_2; 1, 0) \sim (p_1 \text{ or } p_2; \beta, \beta)$, by Prize Monotonicity, we have

$$\beta = (1 - \varepsilon)[\alpha p_2 + (1 - \alpha)p_1] + \varepsilon \cdot 0.5.$$

Hence,

$$(p_1 \text{ or } p_2; v_1, v_2) \sim ((1 - \varepsilon)[\alpha p_2 + (1 - \alpha)p_1] + \varepsilon \cdot 0.5; v_1, v_2).$$

For $v_1 < v_2$, by Symmetric Events,

$$(p_1 \text{ or } p_2; v_1, v_2) \sim (1 - p_1 \text{ or } 1 - p_2; v_2, v_1) \sim ((1 - \varepsilon)[\alpha(1 - p_2) + (1 - \alpha)(1 - p_1)] + \varepsilon \cdot 0.5; v_2, v_1).$$

By Simple vNM, the preference over uncertain acts can be represented as stated in the Theorem.

E Theories based on the smooth preference

In this appendix, I consider alternative models of updating with uncertain information and uncertain priors that are based on the smooth model (Klibanoff et al., 2005). In a decision problem without updating in my experiment, the evaluation of a bet with uncertain prior (p_h or p_l) under the smooth model is

$$\phi^{-1}\left(0.5\phi(p_h) + 0.5\phi(p_l)\right)$$

where $\phi(\cdot)$ is an increasing function. For each prior p, the agent evaluates the bet using the standard von Neumann-Morgenstern (vNM) expected utility. Then the agent assigns each vNM utility index a second-order utility $\phi(\cdot)$, and then she aggregates the second-order utilities by taking the expectation. (The ϕ^{-1} is for the purpose of normalization.) The second-order utility function $\phi(\cdot)$ summarizes an agent's attitude toward uncertainty. If ϕ is concave, then the agent behaves pessimistically in face of uncertainty. The opposite is true if ϕ is convex. The smooth model coincides with standard EU when ϕ is linear. The maxmin model (Gilboa and Schmeidler, 1989) is the limit case when ϕ becomes infinitely concave. However, the smooth model cannot capture uncertainty-induced insensitivity. Hence, I will only discuss the implications of the concavity of ϕ in belief updating problems. I now define a comparative notion of concavity which I will use in the rest of this section: function ϕ_2 is more concave than function ϕ_1 if $\phi_2(\phi_1^{-1}(\cdot))$ is concave.

E.1 Recursive smooth preferences

In an uncertain information problem where the prior is p and the information accuracy is either ψ_h or ψ_l, an agent with a recursive smooth preference (Klibanoff et al., 2009) applies Bayes's rule to calculate the posterior distribution over the outcomes given each possible level of accuracy, $Pr^{Bayes}(G|p, m, \psi_i)$, as well as the updated distribution over levels of accuracy, $Pr^{Bayes}(\psi_i|p, m)$.[46] Then, to obtain the evaluation of the bet conditional on message m, the agent calculates the vNM expected utility for each posterior and aggregate them up by taking the expectation over the second-order utility function ϕ using the updated distribution over levels of accuracy:

$$\phi^{-1}\left(Pr^{Bayes}(\psi_h|p, m)\cdot\phi\left(Pr^{Bayes}(G|p, m, \psi_h)\right) + Pr^{Bayes}(\psi_l|p, m)\cdot\phi\left(Pr^{Bayes}(G|p, m, \psi_l)\right)\right).$$

If ϕ is linear, then the evaluation of the bet after observing message m equals $Pr^{Bayes}(G|p, m, (\psi_h + \psi_l)/2)$, same as the conditional evaluation when the accuracy of information equals $(\psi_1 + \psi_2)/2$ with certainty. In the extreme case of maxmin preferences, i.e. when ϕ becomes infinitely concave, conditional smooth preference is equivalent to Full Bayesian updating (Pires, 2002).

Similarly, in an uncertain prior problem, the agent calculates the Bayesian posterior given each possible prior and aggregate them up using the second-order utility function and the updated likelihood of each prior.

The following proposition summarizes the properties of conditional evaluations under recursive smooth preferences.

Proposition 9 1. *If an agent has a recursive smooth preference and her second-order utility function ϕ is linear, her conditional evaluations of a bet in an uncertain information problem are the same as when the accuracy level of the information is $(\psi_h + \psi_l)/2$ with certainty, and her conditional evaluations of a bet in an uncertain prior problem are the same as when the prior is $(p_h + p_l)/2$ with certainty.*

2. *As ϕ becomes more concave, conditional evaluations of the bet decrease.*

Proof I only prove the second part of the proposition for uncertain information problems. If ϕ_1 is more compound-averse than ϕ_2, then $\phi_1(\phi_2^{-1})$ is strictly concave. Hence, by Jensen's Inequality,

$$\phi_1\left(\phi_2^{-1}\left(\sum_{i=h,l}\phi_2\left(Pr^{Bayes}(G|p, m, \psi_i)Pr^{Bayes}(\psi_i|p, m)\right)\right)\right) \geq \sum_{i=h,l}\phi_1\left(Pr^{Bayes}(G|p, m, \psi_i)Pr^{Bayes}(\psi_i|p, m)\right),$$

$$\Longrightarrow \phi_2^{-1}\left(\sum_{i=h,l}\phi_2\left(Pr^{Bayes}(G|p, m, \psi_i)Pr^{Bayes}(\psi_i|p, m)\right)\right) \geq \phi_1^{-1}\left(\sum_{i=h,l}\phi_1\left(Pr^{Bayes}(G|p, m, \psi_i)Pr^{Bayes}(\psi_i|p, m)\right)\right).$$

[46] $Pr^{Bayes}(\psi_h|p, g) = \frac{p\psi_h+(1-p)(1-\psi_h)}{p\psi_h+(1-p)(1-\psi_h)+p\psi_l+(1-p)(1-\psi_l)}$, $Pr^{Bayes}(\psi_h|p, b) = \frac{p(1-\psi_h)+(1-p)\psi_h}{p(1-\psi_h)+(1-p)\psi_h+p(1-\psi_l)+(1-p)\psi_l}$.

E.2 Dynamically consistent updating

Under Dynamically consistent updating, the agent forms a contingent plan of actions before the message realizes and executes the plan resolutely after observing the message. In our example where the agent chooses between a bet and a sure amount of utils, the contingent plan, $(a(g), a(b))$, specifies an action $a(m) \in \{Bet, Sure\}$ conditional on each of the two possible messages. The optimal plan maximizes utility from the ex-ante perspective. In an uncertain information problem, the agent with a smooth preference optimizes her plan by solving the following problem:

$$\max_{(a(g),a(b))} \phi^{-1}\left(\sum_{i=h,l} .5 \cdot \phi\left(p \cdot (\psi_i U(a(g), G) + (1-\psi_i)U(a(b), G)) + (1-p) \cdot (\psi_i U(a(b), B) + (1-\psi_i)U(a(g), B))\right)\right),$$

where $U(a, s)$ is the util that the agent receive if the outcome of the bet is s and she takes action a. It's straightforward that the payoff of (Bet, Bet) equals p and that of $(Sure, Sure)$ equals the sure amount, both of which independent from the uncertainty in information accuracy.

If ϕ is linear, then the optimal contingent plan is characterized by two thresholds. If the sure amount of util u is no smaller than $Pr^{Bayes}(G|g, (\psi_h + \psi_l)/2)$, then $(Sure, Sure)$ is optimal. If $u \in \left[Pr^{Bayes}(G|g, (\psi_h + \psi_l)/2), Pr^{Bayes}(G|b, (\psi_h + \psi_l)/2)\right]$, then $(Sure, Bet)$ is optimal. If u is no more than $Pr^{Bayes}(G|b, (\psi_h + \psi_l)/2)$, then (Bet, Bet) is optimal. Note that the actions prescribed by the optimal contingent plan coincide with the ex-post optimal action under the recursive smooth preference. This shows that with linear second-order utility function and Bayesian updating, there is no conflict between the ex-ante preference and the ex-post preference.

Now we consider the case that ϕ is concave. For each sure amount of utils u, an agent using Dynamically consistent updating chooses the plan of action that leads to the highest ex-ante utility among $(Sure, Sure)$, $(Bet, Sure)$, $(Sure, Bet)$ and (Bet, Bet). It's straightforward to show that $(Sure, Bet)$ can never be optimal, so we are left with the other three plans. The comparison between $(Sure, Sure)$ and (Bet, Bet) is simple. The former generates higher utility than the latter if and only if $u > p$. To compare these two plans with $(Bet, Sure)$, we first invoke Jensen's Inequality to obtain an upper bound for the utility of $(Bet, Sure)$.

$$\phi^{-1}\left(\sum_{i=h,l} 0.5 \cdot \phi\left(p \cdot (\psi_i + (1-\psi_i)u) + (1-p) \cdot \psi_i u\right)\right) < \sum_{i=h,l} 0.5 \cdot \left(p \cdot (\psi_i + (1-\psi_i)u) + (1-p) \cdot \psi_i u\right). \tag{6}$$

For any $u \leq Pr^{Bayes}(G|b, (\psi_h + \psi_l)/2)$, the right-hand side of (6) is smaller than p. This implies that for this range of u, (Bet, Bet) is optimal. For any $u \geq Pr^{Bayes}(G|g, (\psi_h + \psi_l)/2)$, the right-hand side of (6) is smaller than u. This implies that for this range of u, $(Sure, Sure)$ is optimal. Hence, if

$(Bet, Sure)$ is ever optimal, it has to be in the region

$$u \in \left(Pr^{Bayes}(G|g, (\psi_h + \psi_l)/2), Pr^{Bayes}(G|b, (\psi_h + \psi_l)/2) \right).$$

Remember that for an agent with linear second-order utility function, $(Bet, Sure)$ is optimal exactly in the closed interval $u \in \left[Pr^{Bayes}(G|g, (\psi_h + \psi_l)/2), Pr^{Bayes}(G|b, (\psi_h + \psi_l)/2) \right]$. This demonstrates that an pessimistic agent using Dynamically consistent updating is averse to conditioning her action on the realized message because doing so exposes her ex-ante utility to the uncertainty in the information accuracy.

If $\psi_l < 0.5 < \psi_h$ and $(\psi_h + \psi_l)/2$ is sufficiently close (but not necessarily equal) to 0.5, then a pessimistic agent using Dynamically consistent updating may never condition her action on the realized message. In other words, the agent may never find $(Bet, Sure)$ optimal, whatever is the sure amount. To see this, note that if $\psi_h + \psi_l = 1$, then the interval $\left(Pr^{Bayes}(G|b, (\psi_h + \psi_l)/2), Pr^{Bayes}(G|g, (\psi_h + \psi_l)/2) \right)$ is empty. So even if we perturb ψ_h and ψ_l so that they don't add up to 1 exactly, the interval is still empty, which means that $(Bet, Sure)$ is never optimal.

When $(Bet, Sure)$ is optimal under some values of u, denote the largest such value by $u(g)$ and the smallest by $u(b)$. I interpret these two values as the evaluations of the bet conditional on message g and b, respectively. Now I show that $u(g)$ decreases and $u(b)$ increases as the second-order utility function becomes more concave.

If ϕ_1 is more concave than ϕ_2, then $\phi_1(\phi_2^{-1})$ is strictly concave. Hence, for all u such that

$$\phi_2(u) \geq \sum_{i=h,l} 0.5 \cdot \phi_2 \left(p \cdot (\psi_i + (1 - \psi_i)u) + (1 - p) \cdot \psi_i u \right),$$

we have

$$\begin{aligned} \phi_1(u) &\geq \phi_1 \left(\phi_2^{-1} \left(\sum_{i=h,l} 0.5 \cdot \phi_2 \left(p \cdot (\psi_i + (1 - \psi_i)u) + (1 - p) \cdot \psi_i u \right) \right) \right) \\ &> \sum_{i=h,l} 0.5 \cdot \phi_1 \left(p \cdot (\psi_i + (1 - \psi_i)u) + (1 - p) \cdot \psi_i u \right). \end{aligned}$$

The second inequality invokes Jensen's Inequality. Denote by $u_1(g)$ and $u_2(g)$ the largest values of u such that $(Bet, Sure)$ is optimal when the second-order utility function is ϕ_1 and ϕ_2, respectively. Then we have $u_1(g) < u_2(g)$. We can similarly define $u_1(b)$ and $u_2(b)$ and prove that $u_1(b) < u_2(b)$.

In a problem with uncertain prior (p_h or p_l) and simple information with accuracy ψ, an agent with smooth preference who uses Dynamically consistent updating solves the following problem:

$$\max_{(a(g),a(b))} \phi^{-1}\left(\sum_{i=h,l} 0.5\cdot\phi\left(p_i\cdot(\psi U(a(g),G)+(1-\psi)U(a(b),G))+(1-p_i)\cdot(\psi U(a(b),B)+(1-\psi)U(a(g),B))\right)\right).$$

The ex-ante payoff of the plan (Bet, Bet) is $\phi^{-1}(0.5\phi(p_h)+0.5\phi(p_l))$ and that of $(Sure, Sure)$ is the sure amount of utils u. Similar as before, $(Sure, Bet)$ is never optimal, so we only consider the other three plans. Suppose that the second-order utility function is strictly concave. Plugging in $u = Pr^{Bayes}(G|(p_h+p_l)/2, g, \psi)$ to the ex-ante payoff of $(Sure, Bet)$ and applying Jensen's Inequality, we get

$$\phi^{-1}\left(\sum_{i=h,l} 0.5\cdot\phi\left(p_i\cdot(\psi+(1-\psi)u)+(1-p_i)\cdot\psi u\right)\right) < \sum_{i=h,l} 0.5\cdot(p_i\cdot(\psi+(1-\psi)u)+(1-p_i)\cdot\psi u) = u.$$

This implies that for an uncertainty-averse agent, the threshold between $(Sure, Sure)$ and $(Bet, Sure)$ is lower than $Pr^{Bayes}(G|(p_h+p_l)/2, g, \psi)$. Moreover, it can be shown that this threshold becomes lower as ϕ becomes more concave.

Similarly, plugging in $u = Pr^{Bayes}(G|(p_h+p_l)/2, b, \psi)$ and applying Jensen's Inequality, we obtain that

$$\phi^{-1}\left(\sum_{i=h,l} 0.5\cdot\phi\left(p_i\cdot(\psi+(1-\psi)u)+(1-p_i)\cdot\psi u\right)\right) > \phi^{-1}(0.5\phi(p_h)+0.5\phi(p_l)).$$

This implies that the threshold between $(Bet, Sure)$ and (Bet, Bet) is lower than $Pr^{Bayes}(G|(p_h+p_l)/2, b, \psi)$.

It can also be shown that as ϕ becomes more concave, the threshold between $(Sure, Sure)$ and $(Bet, Sure)$, $u(g)$, decreases. If ϕ_1 is more concave than ϕ_2, then $\phi_1(\phi_2^{-1})$ is strictly concave. Hence, for all u such that

$$\phi_2(u) \geq \sum_{i=h,l} 0.5\cdot\phi_2\left(p_i\cdot(\psi+(1-\psi)u)+(1-p_i)\cdot\psi u\right),$$

we have

$$\begin{aligned}\phi_1(u) &\geq \phi_1\left(\phi_2^{-1}\left(\sum_{i=h,l} 0.5\cdot\phi_2\left(p_i\cdot(\psi+(1-\psi)u)+(1-p_i)\cdot\psi u\right)\right)\right)\\ &> \sum_{i=h,l} 0.5\cdot\phi_1\left(p_i\cdot(\psi+(1-\psi)u)+(1-p_i)\cdot\psi u\right).\end{aligned}$$

The second inequality invokes Jensen's Inequality. Denote by $u_1(g)$ and $u_2(g)$ the largest values of u such that $(Bet, Sure)$ is optimal when the second-order utility function is ϕ_1 and ϕ_2, respectively. Then we have $u_1(g) < u_2(g)$.

The following proposition summarizes the results above.

Proposition 10
1. *If an agent with smooth preference uses Dynamically consistent updating and her second-order utility function ϕ is linear, then her conditional evaluations of a bet in an uncertain information problem are the same as when the information accuracy is $(\psi_h + \psi_l)/2$ with certainty. Her conditional evaluations of a bet in an uncertain prior problem are also identical to the case where the prior is $(p_h + p_l)/2$ with certainty.*

2. *In an uncertain information problem under Dynamically consistent updating, as the second-order utility function $\phi(\cdot)$ becomes more concave, the evaluation conditional on good news, $u(g)$, decreases and that conditional on bad news, $u(b)$, increases.*

3. *In an uncertain prior problem under Dynamically consistent updating, if the second-order utility function is concave, then $u(g) < Pr^{Bayes}(G|(p_h+p_l)/2, g, \psi)$ and $u(b) < Pr^{Bayes}(G|(p_h+p_l)/2, b, \psi)$. As the second-order utility function $\phi(\cdot)$ becomes more concave, $u(g)$ decreases.*

E.3 Maximum likelihood updating

Similar as with ε-α-maxmin preferences, attitudes toward uncertainty generically do not affect the conditional evaluations under Maximum likelihood updating when only one accuracy level or one prior has the highest likelihood given the realized message. When there is a tie, Maximum likelihood updating coincides with the recursive smooth preference.

E.4 Pre-screening

Similar to Maximum likelihood updating, in uncertain information problems, Pre-screening (Cheng and Hsiaw, 2018) also puts excessive weights on more likely levels of accuracy conditional on the message, but it is less extreme in doing so. Specifically, a Pre-screener's posterior distribution over levels of accuracy conditional on the realized message is given by

$$Pr^{PS}(\psi_h|p,g) = \frac{Pr^{Bayes}(\psi_h|p,g)\cdot p\psi_h + (1 - Pr^{Bayes}(\psi_h|p,g))(1-p)(1-\psi_h)}{\sum_{i=h,l}\left(Pr^{Bayes}(\psi_i|p,g)\cdot p\psi_i + (1 - Pr^{Bayes}(\psi_i|p,g))(1-p)(1-\psi_i)\right)}$$

and

$$Pr^{PS}(\psi_h|p,b) = \frac{Pr^{Bayes}(\psi_h|p,b)\cdot p\psi_h + (1-Pr^{Bayes}(\psi_h|p,b))(1-p)(1-\psi_h)}{\sum_{i=h,l}(Pr^{Bayes}(\psi_i|p,b)\cdot p\psi_i + (1-Pr^{Bayes}(\psi_i|p,b))(1-p)(1-\psi_i))}.$$

Intuitively, a Pre-screener uses the realized message twice when updating her belief over the accuracy levels. This leads to confirmation bias. If $p > 0.5$, then $Pr^{PS}(\psi_h|p,g) > Pr^{Bayes}(\psi_h|p,g)$ and $Pr^{PS}(\psi_h|p,b) < Pr^{Bayes}(\psi_h|p,b)$; if, instead, $p < 0.5$, then $Pr^{PS}(\psi_h|p,g) < Pr^{Bayes}(\psi_h|p,g)$ and $Pr^{PS}(\psi_h|p,b) > Pr^{Bayes}(\psi_h|p,b)$. A pre-screener's posterior beliefs over the accuracy levels coincide with a Bayesian agent's when $p = 0.5$.

Same as in recursive smooth preferences and Maximum likelihood updating, a Pre-screener's posterior distribution over the outcomes given each possible level of accuracy is equal to the correct Bayesian posterior $Pr^{Bayes}(\cdot|p,m,\psi_i)$. Then the Pre-screener's evaluation of the bet conditional on message m is given by

$$\phi^{-1}\left(Pr^{PS}(\psi_h|p,m)\cdot\phi\left(Pr^{Bayes}(G|p,m,\psi_h)\right) + Pr^{PS}(\psi_l|p,m)\cdot\phi\left(Pr^{Bayes}(G|p,m,\psi_l)\right)\right).$$

Combining Pre-screening with smooth preferences sometimes leads to ambiguous implications on the conditional evaluations of bets because confirmation bias and pessimism/optimism may have opposite effects. The following proposition summarizes the unambiguous implications.

Proposition 11

1. *If $p < 0.5$ and the second-order utility function ϕ is concave, then a Pre-screener's conditional evaluations of the bet are lower than the case where the accuracy level of the information is known to be $(\psi_h + \psi_l)/2$.*

2. *If $p > 0.5$ and the second-order utility function ϕ is convex, then a Pre-screener's conditional evaluations of the bet are higher than the case where the accuracy level of the information is known to be $(\psi_h + \psi_l)/2$.*

3. *If $p = 0.5$, then Pre-screening coincides with the recursive smooth preference.*

4. *As ϕ becomes more concave, the conditional evaluations of the bet become lower.*

F Segal's two-stage model

Segal (1987, 1990) proposes a two-stage model of uncertain bets. In a decision problem without updating, the evaluation of a bet with simple prior p is $f(p)$, where $f : [0,1] \to [0,1]$ is an increasing function which satisfies $f(0) = 0$ and $f(1) = 1$. The function f is interpreted as a probability weighting function. The utility of a bet with uncertain prior (p_h or p_l) under the Segal's model is

$$f(0.5)f(p_h) + (1 - f(0.5))f(p_l).$$

The agent calculates the anticipated utility given each prior: $f(p_h)$ and $f(p_l)$. Then she aggregates the anticipated utilities given each prior by applying the same anticipated utility on the second stage.

In Segal's model, the probability weighting function f summarizes the attitude toward uncertainty. Hence, I investigate what properties of f can generate the two-fold pattern of uncertainty-induced aversion and insensitivity in problems with uncertain priors without additional information. Consider two bets. The first one's odds is either 90% or 50%, and the second one's odds is either 50% or 10%. In my experiment, the CE of the first bet is typically lower than its simple counterpart ($p = 70\%$) and the CE of the second bet is about the same as that of a 30%-odds simple bet. This implies that the modal preference satisfies

$$f(0.5)f(0.9) + (1 - f(0.5))f(0.5) < f(0.7)$$

and

$$f(0.5)f(0.5) + (1 - f(0.5))f(0.1) = f(0.3).$$

If $f(0.5) \geq 0.5$, then $f(0.7) > 0.5(f(0.9) + f(0.5))$; if $f(0.5) \leq 0.5$, then $f(0.3) \leq 0.5(f(0.5) + f(0.1))$. This appears at odds with the typically observed inverse S-shaped probability weighting function.

One natural way to incorporate belief updating into Segal's model is to apply Bayes' rule and assume that the probability weighting function f is unaffected by information.[47] Specifically, I assume that the utility of a simple bet with prior p conditional on information whose accuracy is either ψ_h or ψ_l is

$$f\left(Pr^{Bayes}(\psi_h|p,g)\right) f\left(Pr^{Bayes}(G|p,g,\psi_h)\right) + \left(1 - f\left(Pr^{Bayes}(\psi_h|p,g)\right)\right) f\left(Pr^{Bayes}(G|p,g,\psi_l)\right)$$

for good news and

$$f\left(Pr^{Bayes}(\psi_l|p,b)\right) f\left(Pr^{Bayes}(G|p,b,\psi_l)\right) + \left(1 - f\left(Pr^{Bayes}(\psi_l|p,b)\right)\right) f\left(Pr^{Bayes}(G|p,b,\psi_h)\right)$$

for bad news. The conditional valuation when the information accuracy is $(\psi_h+\psi_l)/2$ is $f\left(Pr^{Bayes}(G|p,m,(\psi_h+\psi_l)/2)\right)$, $m \in \{g,b\}$. If f is convex, then by Jensen's inequality, the evaluations conditional on uncertain

[47] I am unaware of any formal treatment of belief updating under probability weighting.

information is lower than those conditional on the corresponding simple information. If f is concave, then the reverse is true. However, it's unclear how insensitivity is related to the comparison.

Similarly, the utility of a bet with uncertain prior conditional on simple message m is

$$f\left(Pr^{Bayes}(p_h|m,\psi)\right) f\left(Pr^{Bayes}(G|p_h,m,\psi)\right)+\left(1-f\left(Pr^{Bayes}(p_h|m,\psi)\right)\right) f\left(Pr^{Bayes}(G|p_l,m,\psi)\right).$$

Jensen's inequality can be similarly applied to derive the relation between the convexity of f and pessimism.

G Additional results on the correlation between different kinds of uncertainty attitudes

In this section, I derive tests of correlations between attitudes toward different kinds of uncertainty for each model considered in this paper. The tests are based on equations between signs of premiums, which are summarized in Table G.1. For example, the left- and upper-most cell states that if an agent's attitudes toward uncertain priors (in problems without updating) and uncertain information can be described by the same ε-α-maxmin preference and the agent uses Full Bayesian updating adapted to the generalized Bayes' rule, then for all $p \in (0,1)$, $0 < x < 1$ and $0 < y < 1$ such that $x + y \geq 1$, the following two equations hold: $SP(p, g, x \text{ or } y) = SP(x \text{ or } y)$ and $SP(p, b, x \text{ or } y) = SP(1-y \text{ or } 1-x)$. So to test that subjects' attitudes toward uncertainty in priors (in problems without updating) and uncertainty in information accuracy are correlated under the assumptions of ε-α-maxmin preference and the adapted Full Bayesian updating, I compute the correlation coefficients between $SP(p, g, x \text{ or } y)$ and $SP(x \text{ or } y)$ and between $SP(p, b, x \text{ or } y)$ and $SP(1-y \text{ or } 1-x)$.

In the rest of this section, I will prove the equations in Table G.1 for each model.

	Uncertain priors w/o updating and uncertain info	Uncertain priors w/o updating and uncertain priors with updating	Uncertain priors with updating and uncertain info
FB	(If $x + y \geq 1$) $SP(p, g, x \text{ or } y) = SP(x \text{ or } y)$ $SP(p, b, x \text{ or } y) = SP(1 - y \text{ or } 1 - x)$	$SP(x \text{ or } y, m, \psi) = SP(x \text{ or } y)$	(If $x + y \geq 1$) $SP(p, g, x \text{ or } y) = SP(x \text{ or } y, m, \psi)$ $SP(p, b, x \text{ or } y) = SP(1 - y \text{ or } 1 - x, m, \psi)$
DC	(If $x + y > 1$) $SP(p, g, x \text{ or } y) = SP(x \text{ or } y)$ $SP(p, b, x \text{ or } y) = -SP(x \text{ or } y)$	$SP(x \text{ or } y, m, \psi) = SP(x \text{ or } y)$	(If $x + y > 1$) $SP(p, g, x \text{ or } y) = SP(x \text{ or } y, m, \psi)$ $SP(p, b, x \text{ or } y) = -SP(x \text{ or } y, m, \psi)$
ML	$SP(50\%, g, 90\% \text{ or } 50\%) = SP(90\% \text{ or } 50\%)$ $SP(50\%, b, 90\% \text{ or } 50\%) = SP(10\% \text{ or } 50\%)$	$SP(90\% \text{ or } 50\%, -, 50\%) = SP(90\% \text{ or } 50\%)$ $SP(10\% \text{ or } 50\%, -, 50\%) = SP(10\% \text{ or } 50\%)$	$SP(50\%, g, 90\% \text{ or } 50\%)$ $= SP(90\% \text{ or } 50\%, -, 50\%)$ $SP(50\%, b, 90\% \text{ or } 50\%)$ $= SP(10\% \text{ or } 50\%, -, 50\%)$
R-S	$SP(50\%, g, 90\% \text{ or } 50\%) = SP(90\% \text{ or } 50\%)$ $SP(50\%, b, 90\% \text{ or } 50\%) = SP(10\% \text{ or } 50\%)$	$SP(90\% \text{ or } 50\%, -, 50\%) = SP(90\% \text{ or } 50\%)$ $SP(10\% \text{ or } 50\%, -, 50\%) = SP(10\% \text{ or } 50\%)$	(If $x + y \geq 1$ and $z \geq 50\%$) $SP(z, g, x \text{ or } y) = SP(x \text{ or } y, g, z)$ $SP(z, b, x \text{ or } y) = SP(1 - y \text{ or } 1 - x, g, z)$ $SP(1 - z, g, x \text{ or } y) = SP(x \text{ or } y, b, z)$ $SP(1 - z, b, x \text{ or } y) = SP(1 - y \text{ or } 1 - x, b, z)$
DC-S	-	-	(If $x + y > 1$ and $z \geq 50\%$) $SP(z, g, x \text{ or } y) = SP(x \text{ or } y, g, z)$
ML-S	$SP(50\%, g, 90\% \text{ or } 50\%) = SP(90\% \text{ or } 50\%)$ $SP(50\%, b, 90\% \text{ or } 50\%) = SP(10\% \text{ or } 50\%)$	$SP(90\% \text{ or } 50\%, -, 50\%) = SP(90\% \text{ or } 50\%)$ $SP(10\% \text{ or } 50\%, -, 50\%) = SP(10\% \text{ or } 50\%)$	$SP(50\%, g, 90\% \text{ or } 50\%)$ $= SP(90\% \text{ or } 50\%, -, 50\%)$ $SP(50\%, b, 90\% \text{ or } 50\%)$ $= SP(10\% \text{ or } 50\%, -, 50\%)$
PS-S	$SP(50\%, g, 90\% \text{ or } 50\%) = SP(90\% \text{ or } 50\%)$ $SP(50\%, b, 90\% \text{ or } 50\%) = SP(10\% \text{ or } 50\%)$	-	-

Table G.1: Summary of theoretical predictions on the relations between signs of premiums of different kinds of uncertainty

Notes: This table summarizes the relations between the signs of premiums of different kinds of uncertainty as predicted by several theories if attitudes toward different kinds of uncertainty are described by the same ε-α-maxmin preference or the same smooth preference. “FB": Full Bayesian updating adapted to the generalized Bayes’ rule coupled with ε-α-maxmin preferences. “DC": Dynamically consistent updating adapted to the generalized Bayes’ rule coupled with ε-α-maxmin preferences. “ML": Maximum likelihood updating coupled with ε-α-maxmin preferences. “R-S": Recursive smooth preference. “DC-S": Smooth preference with Dynamically consistent updating. “ML-S": Smooth preference with Maximum likelihood updating. “PS-S": Smooth preference with pre-screening. If no restriction is mentioned, p, ψ, x, y, z are arbitrary numbers in the interval $(0, 1)$, and m is an arbitrary message (g or b). Equations under “Uncertain priors w/o updating and uncertain info" are equality relations predicted by the corresponding theory if the attitudes toward uncertainty in priors (without information) and the attitudes toward uncertainty in information are described by the same preference. The other two columns are analogous.

G.1 ε-α-maxmin preferences

Suppose an agent's attitudes toward uncertain priors (in problems without updating) is described by an ε-α-maxmin preference, then her CE of an uncertain bet is given by

$$CE(p_h \text{ or } p_l) = M\left(W(p_h, p_l; \varepsilon, \alpha)\right),$$

where $M : [0,1] \to \mathbb{R}_+$ is an increasing function that maps the (subjective) winning odds of a bet to its CE. Suppose that the same ε-α-maxmin preference also describes her attitudes toward uncertainty in information accuracy and that she follows the Full Bayesian updating rule adapted to the generalized Bayes' rule. Then the agent's CE for a simple bet is

$$CE(p, g, \psi_h \text{ or } \psi_l) = M\left(Pr^{GB}(G|p, g, W(\psi_h, \psi_l; \varepsilon, \alpha))\right)$$

conditional on uncertain good news and

$$CE(p, b, \psi_h \text{ or } \psi_l) = M\left(Pr^{GB}(G|p, b, W(\psi_l, \psi_h; \varepsilon, \alpha))\right)$$

conditional on uncertain bad news. Note that $M(\cdot)$ is increasing and the generalized Bayesian posterior is increasing in information accuracy conditional on good news and decreasing conditional on bad news. This implies that if $0 \leq y < x \leq 1$ and $x + y \geq 1$, then for any p,

$$\begin{aligned} SP(p, g, x \text{ or } y) &= sign\left(M\left(Pr^{GB}(G|p, g, \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, g, W(x, y; \varepsilon, \alpha))\right)\right) \\ &= sign\left(M\left(\frac{x+y}{2}\right) - M\left(W(x, y; \varepsilon, \alpha)\right)\right) \\ &= SP(x \text{ or } y) \end{aligned} \tag{7}$$

and

$$\begin{aligned} SP(p, g, x \text{ or } y) &= sign\left(M\left(Pr^{GB}(G|p, b, \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, b, W(y, x; \varepsilon, \alpha))\right)\right) \\ &= sign\left(M\left(Pr^{GB}(G|p, g, 1 - \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, g, 1 - W(y, x; \varepsilon, \alpha))\right)\right) \\ &= sign\left(M\left(\frac{1-y+1-x}{2}\right) - M\left(W(1-y, 1-x; \varepsilon, \alpha)\right)\right) \\ &= SP(1-y \text{ or } 1-x). \end{aligned} \tag{8}$$

Suppose, instead, that the agent uses Dynamically consistent updating adapted to the generalized Bayes' rule. Then the hypothesis that the same ε-α-maxmin preference applies to both uncertain priors (in problems without updating) and uncertain information implies that the CE of a simple bet conditional on uncertain information is

$$CE(p, m, \psi_h \text{ or } \psi_l) = M\left(Pr^{GB}(G|p, m, \max\{W(\psi_h, \psi_l; \varepsilon, \alpha), 0.5\})\right),$$

which in turn implies that if $0 < p < 1$, $0 \leq y < x \leq 1$ and $x + y > 1$, then

$$\begin{aligned} SP(p, g, x \text{ or } y) &= sign\left(M\left(Pr^{GB}(G|p, g, \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, g, \max\{W(\psi_h, \psi_l; \varepsilon, \alpha), 0.5\})\right)\right) \\ &= sign\left(M\left(Pr^{GB}(G|p, g, \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, g, W(x, y; \varepsilon, \alpha))\right)\right) \\ &= sign\left(M\left(\frac{x+y}{2}\right) - M\left(W(x, y; \varepsilon, \alpha)\right)\right) \\ &= SP(x \text{ or } y) \end{aligned} \tag{9}$$

and

$$\begin{aligned} SP(p, g, x \text{ or } y) &= sign\left(M\left(Pr^{GB}(G|p, b, \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, b, \max\{W(\psi_h, \psi_l; \varepsilon, \alpha), 0.5\})\right)\right) \\ &= -sign\left(M\left(Pr^{GB}(G|p, g, \frac{x+y}{2})\right) - M\left(Pr^{GB}(G|p, g, \max\{W(\psi_h, \psi_l; \varepsilon, \alpha), 0.5\})\right)\right) \\ &= -SP(x \text{ or } y). \end{aligned} \tag{10}$$

Last, if the agent uses Maximum likelihood updating adapted to the generalized Bayes' rule,[48] then uncertainty attitudes only have a bite on the conditional CEs if the prior is 50%, in which case the prediction coincides with Full Bayesian updating. So in this scenario, Equations (7) and (8) restricted to $p = 50\%$ are the implications of an agent having the same ε-α-maxmin preference toward uncertain priors (in problems without updating) and uncertain information.

Note that the validity of Equations (7) to (10) is independent of the agent's risk preference M and the parameters in the generalized Bayes' rule. The two sides of each equation are also constructed using non-overlapping parts of data. Hence, the correlation between the two sides of each equation constitutes a test of whether subjects' attitudes toward uncertainty in priors (without information) and uncertainty in information accuracy are correlated, *given the theories under which the equation is valid*.

There is one equation that is valid under all three theories and can be the basis of a correlation

[48] The Maximum likelihood updating adapted to the generalized Bayes' rule has the same selection rule as Maximum likelihood updating under Bayes' rule. The difference is that given the selected prior(s)/information accuracy level(s), beliefs are updated using the adapted Full Bayesian updating.

test using data from my experiment:

$$SP(50\%, g, 90\% \text{ or } 50\%) = SP(90\% \text{ or } 50\%). \tag{11}$$

Now I turn to correlations that involve attitudes toward uncertain priors in problems with updating. For an ε-α-maxmin agent who uses the adapted Full Bayesian updating or the adapted Dynamically consistent updating, the CE for an uncertain bet conditional on simple information is

$$CE(p_h \text{ or } p_l, m, \psi) = M\left(Pr^{GB}(G|W(p_h, p_l; \varepsilon, \alpha), m, \psi)\right).$$

Since M is increasing and the generalized Bayesian posterior is increasing in the prior, if an agent uses the same ε-α-maxmin model for uncertainty in priors in problems with and without belief updating, then for any $0.5 \leq \psi < 1$ and $0 \leq y < x \leq 1$,

$$SP(x \text{ or } y, m, \psi) = SP(x \text{ or } y). \tag{12}$$

If the agent uses the adapted Maximum likelihood updating, then Equation (12) is valid if $\psi = 50\%$. Hence, if I require the correlation tests in my experiment to be valid under all three theories, then they need to be based on the equations

$$SP(90\% \text{ or } 50\%, -, 50\%) = SP(90\% \text{ or } 50\%) \tag{13}$$

and

$$SP(10\% \text{ or } 50\%, -, 50\%) = SP(10\% \text{ or } 50\%) \tag{14}$$

Now suppose that an agent uses the same ε-α-maxmin model for uncertainty in information accuracy and uncertainty in priors (in problems with updating). If she uses the adapted Full Bayesian updating, then for any $0 < p < 1$, $0.5 \leq \psi < 1$ and x and y such that $0 \leq y < x \leq 1$ and $x + y \geq 1$,

$$SP(x \text{ or } y, m, \psi) = SP(p, g, x \text{ or } y), \tag{15}$$

and

$$SP(1 - y \text{ or } 1 - x, m, \psi) = SP(p, b, x \text{ or } y). \tag{16}$$

If she uses the adapted Dynamically consistent updating, then for any $0 < p < 1$, $0.5 \leq \psi < 1$ and x and y such that $0 \leq y < x \leq 1$ and $x + y > 1$, Equation (15) holds and

$$SP(x \text{ or } y, m, \psi) = -SP(p, b, x \text{ or } y). \tag{17}$$

If the agent uses the adapted Maximum likelihood updating, then Equations (15) and (16) hold when $p = \psi = 50\%$. So the equation that is valid under all three theories and can form a basis for a correlation test using data in my experiment is

$$SP(90\% \text{ or } 50\%, -, 50\%) = SP(50\%, g, 90\% \text{ or } 50\%). \tag{18}$$

G.2 Smooth preferences

In the context of smooth preferences, the second-order utility function ϕ summarizes an agent's attitude toward uncertainty about an issue. To incorporate issue preference, I allow ϕ to be different for uncertainty in priors (in problems without updating), uncertainty in priors with additional information, and uncertainty in information accuracy. Denote the three second-order utility functions by ϕ_o, ϕ_p and ϕ_i, respectively. If $\phi_o = \phi_i = \phi$ and the agent uses a recursive smooth preference, Maximum likelihood updating or Pre-screening, then the valuation of a simple bet with prior 50% conditional on good news whose accuracy is either 90% or 50% is

$$\begin{aligned}&\phi^{-1}\left(Pr^{Bayes}(90\%|50\%, g) \cdot \phi\left(Pr^{Bayes}(G|50\%, g, 90\%)\right) + Pr^{Bayes}(50\%|50\%, g) \cdot \phi\left(Pr^{Bayes}(G|50\%, g, 50\%)\right)\right) \\ =&\phi^{-1}\left(0.5 \cdot \phi(0.9) + 0.5 \cdot \phi(0.5)\right),\end{aligned}$$

which coincides with the evaluation of an uncertain bet whose odds is either 90% or 50%. If the news on the outcome of the simple bet is bad, then its evaluation would coincide with that of an uncertain bet whose odds is either 10% or 50%. Put in the terms of certainty equivalents, this implies that for agents with smooth preferences who use recursive smooth preferences, Maximum likelihood updating or Pre-screening, the two equations, $CE(50\%, g, 90\% \text{ or } 50\%) = CE(90\% \text{ or } 50\%)$ and $CE(50\%, b, 90\% \text{ or } 50\%) = CE(10\% \text{ or } 50\%)$, hold if the agents have the same attitudes toward uncertainty in priors (in problems without updating) and toward uncertainty in information. Note that under the assumption that the agents are Bayesian when updating with simple priors and simple information, the two equations above imply the corresponding equations on the signs of uncertainty premiums, $SP(50\%, g, 90\% \text{ or } 50\%) = SP(90\% \text{ or } 50\%)$ and $SP(50\%, b, 90\% \text{ or } 50\%) = SP(10\% \text{ or } 50\%)$.

It is analogous to derive the equations that test $\phi_o = \phi_p$: $SP(90\% \text{ or } 50\%, -, 50\%) = SP(90\% \text{ or } 50\%)$ and $SP(10\% \text{ or } 50\%, -, 50\%) = SP(10\% \text{ or } 50\%)$. These two tests are valid if the agents use recursive smooth preferences, Maximum likelihood updating or Pre-screening, and are Bayesian when updating with simple priors and simple information.

To test $\phi_p = \phi_i$, consider an agent who uses a recursive smooth preference. For any $x + y \geq 1$

and $z \geq 50\%$, if $\phi_p = \phi_i = \phi$, then the evaluation of a simple bet with prior z conditional on uncertain good news whose accuracy is either x or y is

$$\begin{aligned}
&\phi^{-1}\begin{pmatrix} Pr^{Bayes}(\psi = x|p = z, g) \cdot \phi\left(Pr^{Bayes}(G|p = z, g, \psi = x)\right) \\ +Pr^{Bayes}(\psi = y|p = z, g) \cdot \phi\left(Pr^{Bayes}(G|p = z, g, \psi = y)\right) \end{pmatrix} \\
=&\phi^{-1}\begin{pmatrix} Pr^{Bayes}(p = x|g, \psi = z) \cdot \phi\left(Pr^{Bayes}(G|p = x, g, \psi = z)\right) \\ +Pr^{Bayes}(p = y|g, \psi = z) \cdot \phi\left(Pr^{Bayes}(G|p = y, g, \psi = z)\right) \end{pmatrix},
\end{aligned}$$

which coincides with the evaluation of an uncertain bet whose prior is either x or y conditional on simple good news whose accuracy is z. Hence, $CE(z, g, x \text{ or } y) = CE(x \text{ or } y, g, z)$. This in turn implies that the test $SP(z, g, x \text{ or } y) = SP(x \text{ or } y, g, z)$ is valid for agents who use recursive smooth preferences. Similarly, $SP(z, b, x \text{ or } y) = SP(1-y \text{ or } 1-x, g, z)$, $SP(1-z, g, x \text{ or } y) = SP(x \text{ or } y, b, z)$ and $SP(1 - z, b, x \text{ or } y) = SP(1 - y \text{ or } 1 - x, b, z)$ are also valid.

For an agent who use Dynamically consistent updating, I consider the sure amount of utils u that makes her indifferent between the contingent plan $(Sure, Sure)$ and $(Bet, Sure)$. The ex-ante utility of the former plan is always u, so I focus on the latter. For any $x + y > 1$ and $z \geq 50\%$, if $\phi_p = \phi_i = \phi$, then the ex-ante utility of $(Bet, Sure)$ is

$$\phi^{-1}\left(\sum_{\psi=x,y} 0.5 \cdot \phi\left(z\psi + z(1-\psi)u + (1-z)\psi u\right)\right)$$

for a simple bet whose prior is z with information whose accuracy is either x or y. This ex-ante utility can be rewritten as

$$\phi^{-1}\left(\sum_{p=x,y} 0.5 \cdot \phi\left(pz + p(1-z)u + (1-p)zu\right)\right)$$

which is the ex-ante utility of $(Bet, Sure)$ for an uncertain bet whose prior is either x or y with simple information whose accuracy is z. This implies that the sure amount of utils u that makes her indifferent between $(Sure, Sure)$ and $(Bet, Sure)$ is the same for the two decision problems. This in turn implies that the equation $SP(z, g, x \text{ or } y) = SP(x \text{ or } y, g, z)$ is valid for an agent who uses Dynamically consistent updating and for whom $\phi_p = \phi_i$ holds.

G.3 Segal's two-stage model

The probability weighting function f determines the uncertainty attitudes in Segal's two-stage model. To incorporate issue preference into Segal's model, I allow for different probability weighting functions for simple risk, uncertainty in priors in problems without updating, uncertainty in priors in problems with updating, and uncertainty in information accuracy. For example, the utility of an uncertain bet is

$$f_o(0.5)f_r(p_h) + (1 - f_o(0.5))f_r(p_l),$$

that of an uncertain bet conditional on simple information is

$$f_p\left(Pr^{Bayes}(p_h|m,\psi)\right) f_r\left(Pr^{Bayes}(G|p_h,m,\psi)\right)+\left(1 - f_p\left(Pr^{Bayes}(p_h|m,\psi)\right)\right) f_r\left(Pr^{Bayes}(G|p_l,m,\psi)\right),$$

and those of a simple bet conditional on uncertain information are

$$f_i\left(Pr^{Bayes}(\psi_h|p,g)\right) f_r\left(Pr^{Bayes}(G|p,g,\psi_h)\right)+\left(1 - f_i\left(Pr^{Bayes}(\psi_h|p,g)\right)\right) f_r\left(Pr^{Bayes}(G|p,g,\psi_l)\right)$$

and

$$f_i\left(Pr^{Bayes}(\psi_l|p,b)\right) f_r\left(Pr^{Bayes}(G|p,b,\psi_l)\right)+\left(1 - f_i\left(Pr^{Bayes}(\psi_l|p,b)\right)\right) f_r\left(Pr^{Bayes}(G|p,b,\psi_h)\right).$$

To test whether the three probability weighting functions, f_o, f_p and f_i, are identical, it is straightforward that the same tests used for recursive smooth preferences apply here. The proofs are omitted.

H Additional results on the correlation between ambiguity and compound attitudes

	$\frac{Amb=Comp}{All}$	$\frac{Amb=Comp\neq Simp}{\neg(Amb=Comp=Simp)}$	$\frac{Simp=Amb}{All}$	$\frac{Simp=Comp}{All}$
Info accuracy	36% (22%)	22% (18%)	30%	30%
Priors (without updating)	39% (23%)	26% (16%)	29%	30%
Priors (with updating)	35% (21%)	23% (16%)	28%	26%

Table H.1: Relation between compound uncertainty and ambiguity

Notes: The first column of this table shows the percentages of cases where corresponding compound and ambiguous CEs are identical. Numbers in parentheses are the maximum of these percentages in 500 simulations where compound and ambiguous CEs are randomly permuted among those that share the same simple counterpart. The second column excludes cases where the corresponding simple, compound, and ambiguous CEs are all the same. The third column shows the proportions of cases where the ambiguous CE is equal to the corresponding simple CE, whereas the last column is analogous for the match between compound CEs and simple CEs.

I Additional results on stock market reactions to earnings forecasts

I.1 A model of asset pricing with uncertain information

In this section, I derive the effects of uncertain information accuracy on stock prices in a simple representative-agent model. The model has three dates, labeled 0, 1, and 2. The representative agent owns a share of a stock, which is a claim to a dividend d whose true amount is revealed at date 2. At date 1, a piece of information m about the dividend is realized. At date 0, the representative agent has a rational expectation about the amount of dividend, which is described by the pdf $F(d)$.

If the agent knows the information structure of m, denoted by $\psi(m|d)$, then her expectation about the dividend conditional on m adheres to Bayes' rule:

$$\mathbb{E}(d|m) = \frac{\int_d d \cdot \psi(m|d) \mathrm{d}F(d)}{\int_d \psi(m|d) \mathrm{d}F(d)}.$$

I now focus on the case where the agent does not know the information structure. For example, the information m may be an earnings forecast issued by an analyst who is unfamiliar to the agent. Suppose that the information structure might either be $\psi_1(m|d)$ or $\psi_2(m|d)$, and the two possibilities are equally likely. Then the Bayesian expectation about the dividend conditional on m should be

$$\mathbb{E}^{Bayes}(d|m) = \frac{\int_d d \cdot (0.5\psi_1(m|d) + 0.5\psi_2(m|d))\, \mathrm{d}F(d)}{\int_d (0.5\psi_1(m|d) + 0.5\psi_2(m|d))\, \mathrm{d}F(d)}.$$

In view of the experimental results in this paper, people may not follow Bayes' rule when the information structure is uncertain. Therefore, adapting the ε-α-maxmin EU preference and Full Bayesian updating to the current setting, I assume that the representative agent's conditional expectation about the dividend is given by

$$\mathbb{E}(d|m) = \frac{\int_d d \cdot \Big((1-\varepsilon)[(1-\alpha)\bar{\psi}(m|d) + \alpha\underline{\psi}(m|d)] + \varepsilon \cdot \psi_0(m)\Big)\, \mathrm{d}F(d)}{\int_d \Big((1-\varepsilon)[(1-\alpha)\bar{\psi}(m|d) + \alpha\underline{\psi}(m|d)] + \varepsilon \cdot \psi_0(m)\Big)\, \mathrm{d}F(d)}.$$

I assume that the pdf $\psi_0(m)$ does not depend on the true dividend d, so it represents an uninformative information structure. Which of the two information structures, ψ_1 and ψ_2, receives the relative weight α depends on which one, when mixed with ψ_0, leads to a more pessimistic expectation:

$$\bar{\psi} = arg \max_{\psi \in \{\psi_1, \psi_2\}} \mathbb{E}(d|m) = \frac{\int_d d \cdot ((1-\varepsilon)\psi(m|d) + \varepsilon \cdot \psi_0(m))\, \mathrm{d}F(d)}{\int_d ((1-\varepsilon)\psi(m|d) + \varepsilon \cdot \psi_0(m))\, \mathrm{d}F(d)}$$

and

$$\underline{\psi} = arg \min_{\psi \in \{\psi_1, \psi_2\}} \mathbb{E}(d|m) = \frac{\int_d d \cdot ((1-\varepsilon)\psi(m|d) + \varepsilon \cdot \psi_0(m))\, \mathrm{d}F(d)}{\int_d ((1-\varepsilon)\psi(m|d) + \varepsilon \cdot \psi_0(m))\, \mathrm{d}F(d)}.$$

Simple algebra lead to a counterpart of Proposition 1 in the stock market setting.

Proposition 12 *Assume that the representative investor has an ε-α-maxmin preference and uses Full Bayesian updating.*

1. *If $\varepsilon = 0$ and $\alpha = 0.5$, then her conditional expectations about the dividend coincide with the Bayesian expectations conditional on simple information with information structure $\frac{\psi_1+\psi_2}{2}$;*

2. *As α increases, the conditional expectations decrease;*

3. *As ε increases, the conditional expectations become closer to the prior expectation $\int_d d\mathrm{d}F(d)$.*

A straightforward corollary of Proposition 12 is that if $\alpha > 0.5$ and $\varepsilon > 0$, then the expectation conditional on good news, i.e. m such that $\mathbb{E}^{Bayes}(d|m) > \int_d d\mathrm{d}F(d)$, is lower than the Bayesian benchmark. This is because both $\alpha > 0.5$ and $\varepsilon > 0$ cause the agent's expectation to deviate from the Bayesian benchmark downwards. For bad news, on the other hand, the comparison with the Bayesian benchmark is ambiguous.

To study the implications on stock prices, I assume for simplicity that the representative agent is risk neutral, does not discount the future, and only cares about the dividend at date 2. Then the stock price at each date is equal to the expectation on that date about the dividend. Also, the abnormal returns at date 2 are hence $R_2 = d - \mathbb{E}(d|m)$. In view of the corollary to Proposition 12, if m is good news, then the abnormal returns are expected to be positive. [49]

I.2 Variable definitions and summary statistics

I.3 Robustness checks for results on stock market reactions to earnings forecasts

[49] Epstein and Schneider (2008) introduce a recursive model where the price at date t is the expectation of the prices at date $t+1$. Making this assumption in my setting would change the stock price at date 0 but not the other results.

	With record			No record		
	N	mean	sd	N	mean	sd
Good news						
Ret[-1,1]	366,050	0.00795	0.0574	31,554	0.00887	0.0717
Ret[-1,22]	365,998	0.0135	0.133	31,553	0.0153	0.164
Ret[-1,43]	365,675	0.0151	0.183	31,539	0.0202	0.231
Ret[-1,64]	364,133	0.0172	0.227	31,462	0.0227	0.293
Ret[-1,EA+1]	364,993	0.0157	0.212	31,403	0.0206	0.272
Bad news						
Ret[-1,1]	562,312	-0.00668	0.0634	46,822	-0.00957	0.0739
Ret[-1,22]	562,235	-0.00770	0.144	46,816	-0.00976	0.164
Ret[-1,43]	561,752	-0.00727	0.194	46,778	-0.00947	0.221
Ret[-1,64]	559,170	-0.00955	0.235	46,653	-0.0128	0.276
Ret[-1,EA+1]	560,105	-0.00994	0.221	46,595	-0.0112	0.272

Table I.1: Returns after forecast revisions

Notes: This table summarizes the size-adjusted returns in different time windows around the forecast announcement, separately for with-record and no-record forecasts and for good news and bad news. It includes only forecasts that meet all the data selection criteria. “EA+1" is the 1st trading day after the announcement of the actual earnings. For the summary statistics of $Ret[-1, EA + 1]$, I exclude observations where the actual earnings announcement happens later than 190 trading days after the forecast announcement. Variable definitions are in Table I.2.

Variable	**Definition**
Main variables	
$Ret[t,T]$	The stock's (buy-hold) returns between the tth and the Tth trading day after the analyst's forecast announcement minus the equal-weighted average returns of stocks in the same size decile in the same period
$NoRecord$	Indicator variable: =0 if the analyst has issued a quarterly earnings forecast on this stock before and the actual earnings of that quarter have been announced; =1 otherwise
$GoodNews$	Indicator variable: =0 if the earnings forecast is a downward revision from the last forecast issued by the same analyst on the same stock's quarterly earnings; =1 if it is an upward revision
Controls	
$ForecastError$	Absolute forecast error is the absolute difference between a forecast and the actual earnings per share, normalized by the stock price two trading days prior to the forecast announcement. Forecast error is absolute forecast error normalized by the standard deviation of absolute forecast errors among all forecasts for the same stock-quarter
$StockExp$/$IndExp$/ $TotExp$	Experience (stock-specific/industry-specific/total): number of days since the analyst's first earnings forecast on the same stock/any stock in the same industry/any stock
$Companies$	Number of stocks covered by the analyst in the same year
$Industries$	Number of industries covered by the analyst in the same year
$Turnover$	Indicator variable: =0 if the analyst has not changed brokerage house in the year; =1 otherwise
$Horizon$	Number of days between the earnings forecast and the end of the forecasted quarter
$DaysElapsed$	Number of days elapsed since the last forecast issued by any analyst on the same firm's quarterly earnings or the firm's last earnings announcement, whichever is later
$BrokerSize$	Number of analysts in the same brokerage house who cover the same stock in the same year
$Coverage$	Number of analysts covering the same firm in the same year
$log(MktCap)$	Logarithm of market capitalization at the end of last year
B/M	Book-to-Market ratio at the end of last year. Winsorized at the 1st and 99th percentiles
$PastReturns$	Size-adjusted returns from seven months before forecast announcement to one month before forecast announcement. Winsorized at the 1st and 99th percentiles
$Volatility$	Standard deviation of the stock's monthly returns in the 24 months before the end of the calendar year prior to the forecast announcement
$Volume$	Average monthly turnover of the stock in the past calendar year

Table I.2: Variable definitions

	With record			No record		
VARIABLES	N	mean	sd	N	mean	sd
GoodNews	943,984	0.394	0.489	81,839	0.403	0.491
ForecastError	937,015	-0.122	0.933	80,787	-0.101	0.941
StockExp	943,984	1,389	1,418	81,839	83.38	242.0
IndExp	943,984	2,379	2,081	81,839	943.2	1,476
TotExp	943,984	3,024	2,351	81,839	1,561	1,920
Companies	943,984	16.73	8.306	81,839	14.01	9.171
Industries	943,984	4.377	2.693	81,839	3.972	2.705
Turnover	943,984	0.0321	0.176	81,839	0.0377	0.191
Horizon	943,984	42.92	46.56	81,839	40.50	50.30
DaysElapsed	943,984	11.49	15.82	81,839	13.24	16.83
BrokerSize	943,984	1.072	0.275	81,839	1.281	0.509
log(MktCap)	943,801	7.826	1.845	81,815	7.151	1.790
B/M	943,784	0.518	0.397	81,815	0.467	0.377
PastReturns	929,349	0.00338	0.297	81,255	0.0455	0.361
Volume	914,191	2.239	1.850	71,545	2.155	1.901
Coverage	943,984	13.73	8.667	81,839	11.39	8.144

Table I.3: Summary statistics

Notes: This table summarizes the indicator variable *GoodNews* and the control variables, separately for with-record and no-record forecasts. It only includes observations that meet all the data selection criteria, i.e. forecast revisions for quarters between January 1st, 1994 and June 30th, 2019 such that on the forecast announcement day, there is neither earnings announcement from the company nor earnings forecast announcements by any other analyst on the same company. Variable definitions are in Table I.2.

	With record			No record		
VARIABLES	N	mean	sd	N	mean	sd
GoodNews	2,412,921	0.393	0.488	168,938	0.398	0.490
ForecastError	2,401,471	-0.161	0.908	167,523	-0.146	0.927
StockExp	2,412,921	1,435	1,447	168,938	75.70	231.7
IndExp	2,412,921	2,470	2,122	168,938	971.5	1,515
TotExp	2,412,921	3,134	2,414	168,938	1,592	1,969
Companies	2,412,921	16.63	7.459	168,938	13.86	8.477
Industries	2,412,921	4.290	2.564	168,938	3.843	2.574
Turnover	2,412,921	0.0262	0.160	168,938	0.0352	0.184
Horizon	2,412,921	49.18	40.83	168,938	46.17	48.98
DaysElapsed	2,412,921	5.349	12.09	168,907	7.603	15.74
BrokerSize	2,412,921	1.070	0.273	168,938	1.287	0.513
log(MktSize)	2,412,502	8.092	1.786	168,890	7.462	1.757
B/M	2,412,450	0.493	0.384	168,889	0.445	0.365
PastReturns	2,375,825	0.00198	0.288	167,802	0.0378	0.353
Volume	2,347,292	2.452	1.912	149,676	2.402	1.996
Coverage	2,412,921	15.92	9.143	168,938	13.61	8.842

Table I.4: Summary statistics (all forecast revisions between 1/1/1994 and 6/30/2019)

Notes: This table summarizes the indicator variable *GoodNews* and the control variables, separately for with-record and no-record forecasts. It includes all forecast revisions for quarters between January 1st, 1994 and June 30th, 2019. Variable definitions are in Table I.2.

	(1) Ret[2,22]	(2) Ret[2,43]	(3) Ret[2,EA+1]
Ret[-1, 1]	0.0179	0.146†	0.252**
	(0.0568)	(0.0775)	(0.0808)
NoRecord	0.000717	0.00170	0.00116
	(0.00120)	(0.00158)	(0.00156)
NoRecord × Ret[-1, 1]	-0.0335	-0.0214	-0.0520
	(0.0227)	(0.0319)	(0.0364)
GoodNews	0.00689***	0.00753***	0.00939***
	(0.000929)	(0.00137)	(0.00149)
GoodNews × Ret[-1, 1]	0.0297	0.0449*	0.0280
	(0.0190)	(0.0222)	(0.0275)
NoRecord × GoodNews	-0.000648	0.000990	0.000416
	(0.00151)	(0.00206)	(0.00210)
NoRecord × GoodNews × Ret[-1, 1]	0.0510	0.100*	0.122*
	(0.0326)	(0.0478)	(0.0496)
Controls	Y	Y	Y
Controls × Ret[-1,1]	Y	Y	Y
Year-Quarter FE	Y	Y	Y
Observations	895740	895168	892678
R^2	0.009	0.012	0.015

Table I.5: Sufficiency of stock market reactions to forecast revisions: different drift lengths

Notes: This table reports the results of Regression (1) with different dependent variables. $Ret[2, 22]$ and $Ret[2, 43]$ are the stock's 1-month and 2-month size-adjusted buy-hold returns starting from the 2nd trading day after the forecast announcement, respectively. "EA+1" is the 1st trading day after the announcement of the actual earnings. In the model $Ret[2, EA + 1]$, I exclude observations where the actual earnings announcement happens later than 190 trading days after the forecast announcement. Three-dimensional (stock, analyst, year-quarter) cluster-robust standard errors in parentheses. $\dagger p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001$

Dependent Var: Ret[2,64]	(1) High innovation	(2) Isolated	(3) After 2004	(4) MktCap>2b	(5) MktCap>10b	(6) Qtr SE&FE
Ret[-1, 1]	0.307*	0.271†	0.369**	-0.00618	-0.625†	
	(0.139)	(0.137)	(0.131)	(0.158)	(0.317)	
NoRecord	0.00213	-0.00211	-0.00210	0.00498†	0.00358	0.000536
	(0.00272)	(0.00229)	(0.00243)	(0.00267)	(0.00281)	(0.00205)
NoRecord × Ret[-1, 1]	-0.0667	-0.0960	-0.105†	-0.0449	0.00603	-0.0264
	(0.0505)	(0.0629)	(0.0570)	(0.0783)	(0.136)	(0.0458)
GoodNews	0.0120***	0.0111***	0.00737***	0.00364*	0.00396*	0.0106***
	(0.00222)	(0.00182)	(0.00185)	(0.00147)	(0.00154)	(0.00176)
GoodNews × Ret[-1 1]	0.0440	0.0715*	0.0163	0.0646*	0.0633	0.0522†
	(0.0339)	(0.0353)	(0.0344)	(0.0264)	(0.0426)	(0.0266)
NoRecord × GoodNews	0.000813	0.00146	-0.000372	-0.00459	-0.00515	0.00131
	(0.00339)	(0.00280)	(0.00287)	(0.00360)	(0.00352)	(0.00247)
NoRecord × GoodNews × Ret[-1, 1]	0.229**	0.153†	0.134†	0.163	0.0522	0.107†
	(0.0780)	(0.0869)	(0.0800)	(0.135)	(0.169)	(0.0620)
Controls	Y	Y	Y	Y	Y	Y
Controls × Ret[-1,1]	Y	Y	Y	Y	Y	Y
Year-Quarter FE	Y	Y	Y	Y	Y	Y
Year-Quarter Slope Effects	N	N	N	N	N	Y
Observations	502879	571449	649583	499077	215867	894004
R^2	0.018	0.013	0.013	0.020	0.017	0.016

Table I.6: Sufficiency of stock market reactions to forecast revisions: robustness checks

Notes: This table reports the results of Regression 1 under different cuts of the data and specifications. “High innovation" restricts attention to forecasts that fall outside the range between the same analyst’s previous forecast and the previous consensus. “Isolated" focuses on observations where in the three-day window centered on the forecast announcement day, there is neither earnings announcement from the company nor forecast announcements by any other analysts on the same company. “After 2004" uses forecasts announced after Jan 1st, 2004. “MktCap>2b" and "MktCap>10b" focus on stocks whose market capitalization is larger than $2 billion and $10 billion, respectively. “Qtr SE&FE" includes the interactions between the Year-Quarter dummies and $Ret[-1, 1]$, in addition to Year-Quarter fixed effects. Three-dimensional (stock, analyst, year-quarter) cluster-robust standard errors in parentheses. $\dagger p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001$

	(1)	(2)
	Ret[-1,1]	Ret[2,64]
Revision	1.077***	2.281*
	(0.258)	(0.972)
NoRecord	-0.00258**	0.000937
	(0.000812)	(0.00297)
NoRecord × Revision	-0.113	-0.566
	(0.179)	(0.480)
GoodNews	0.0116***	0.00641**
	(0.000581)	(0.00196)
GoodNews × Revision	0.438***	2.586***
	(0.118)	(0.727)
NoRecord × GoodNews	0.00448***	-0.00163
	(0.00110)	(0.00397)
NoRecord × GoodNews × Revision	-0.0416	2.822*
	(0.336)	(1.315)
Controls	Y	Y
Controls × Revision	Y	Y
Year-Quarter FE	Y	Y
Observations	503943	502879
R^2	0.026	0.022

Table I.7: Sufficiency of stock market reactions to forecast revisions: magnitudes of revisions

Notes: This table reports the results of the following regression.

$$\begin{aligned} Ret[t,T]_i =& \eta_0 + \eta_1 Revision_i + \eta_2 NoRecord_i + \eta_3 GoodNews_i \\ &+ \eta_4 NoRecord_i \cdot GoodNews_i + \eta_5 Revision_i \cdot GoodNews_i + \eta_6 Revision_i \cdot NoRecord_i \\ &+ \eta_7 Revision_i \cdot NoRecord_i \cdot GoodNews_i + Controls_i + Controls_i \cdot Revision_i + TimeFE_i + \varepsilon_i. \end{aligned} \tag{19}$$

$Revision$ is the difference between an analyst's revised forecast on earnings per share and her previous forecast, normalized by the stock price two trading days prior to the announcement of the revision. I winsorize $Revision$ at the 1st and 99th percentiles. I only include "high innovation" revision, i.e. forecasts that fall outside the range between the same analyst's previous forecast and the previous consensus. Three-dimensional (stock, analyst, year-quarter) cluster-robust standard errors in parentheses. $\dagger p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001$