THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. 138 2023 Issue 4

COGNITIVE UNCERTAINTY*

BENJAMIN ENKE AND THOMAS GRAEBER

This article documents the economic relevance of measuring cognitive uncertainty: people's subjective uncertainty over their ex ante utility-maximizing decision. In a series of experiments on choice under risk, the formation of beliefs, and forecasts of economic variables, we show that cognitive uncertainty predicts various systematic biases in economic decisions. When people are cognitively uncertain—either endogenously or because the problem is designed to be complex—their decisions are heavily attenuated functions of objective probabilities, which gives rise to average behavior that is regressive to an intermediate option. This insight ties together a wide range of empirical regularities in behavioral economics that are typically viewed as distinct phenomena or even as reflecting preferences, including the probability weighting function in choice under risk; base rate insensitivity, conservatism, and sample size effects in belief updating; and predictable overoptimism and -pessimism in forecasts of economic variables. Our results offer a blueprint for how a simple measurement of cognitive uncertainty generates novel insights about what people find complex and how they respond to it. *JEL Code:* D01.

I. INTRODUCTION

This article studies the economic relevance of *cognitive uncertainty*: people's subjective uncertainty over which decision

*We thank the editor and four very constructive referees for helpful comments. We are also grateful to Sebastian Ebert, Thomas Epper, Cary Frydman, Xavier Gabaix, Josh Schwartzstein, Frederik Schwerter, Joel van der Weele, Florian Zimmermann, many job market participants, and seminar audiences. Graeber thanks the Sloan Foundation for postdoctoral funding, and Enke thanks the Foundations of Human Behavior Initiative for financial support.

The Quarterly Journal of Economics (2023), 2021–2067. https://doi.org/10.1093/qje/qjad025.
Advance Access publication on May 27, 2023.

maximizes their expected utility. In the standard economic model, people make decisions that they know may turn out to be ex post suboptimal, but they never exhibit doubts about their ex ante optimality. Similarly, in a large majority of behavioral economics models, people may make systematic mistakes, but they are not nervous that they may be committing errors. Yet introspection and a growing body of psychological evidence discussed below suggest that people often exhibit low confidence in their decisions.

This study proposes that measuring cognitive uncertainty can be productively deployed to predict systematic biases in economic behaviors and to help tie together widely studied behavioral economics anomalies that are typically viewed as distinct phenomena. The main idea consists of two components. (i) Classical anomalies share a common origin, which is that the inherent complexity of economic decisions induces people to make noisy or heuristic decisions instead of solving a problem precisely. These simpler decision modes produce behaviors that are severely attenuated functions of objective problem parameters and are regressive to an intermediate option. (ii) Cognitive uncertainty represents an easily measurable proxy for the unobserved noisiness or heuristic nature of people's decision modes and can thus be used to predict and explain behavior.

We present experiments on decision making under uncertainty: how people reason about probabilities in the valuation of risky lotteries, inference from data, and prediction of future events. As Figure I illustrates using our experimental data, these three literatures have established striking similarities about how objective probabilities map onto people's decisions. First, the left panel depicts the well-known probability weighting function in choice under risk that goes back to Tversky and Kahneman (1992). It illustrates how experimental subjects implicitly treat objective probabilities in choosing between different monetary gambles. Second, the middle panel illustrates the canonical compressed relationship between participants' posterior beliefs and the Bayesian posterior in experimental belief-updating tasks, which shows that people generally overestimate the probability of unlikely events and underestimate the probability of likely ones. Finally, the right panel shows the compressed relationship between respondents' probabilistic estimates and "true" probabilities that has been documented in a wide range of subjective expectations surveys about, for example, stock market returns or inflation rates. The characteristic feature of these three functions is that people's decisions implicitly treat different probabilities

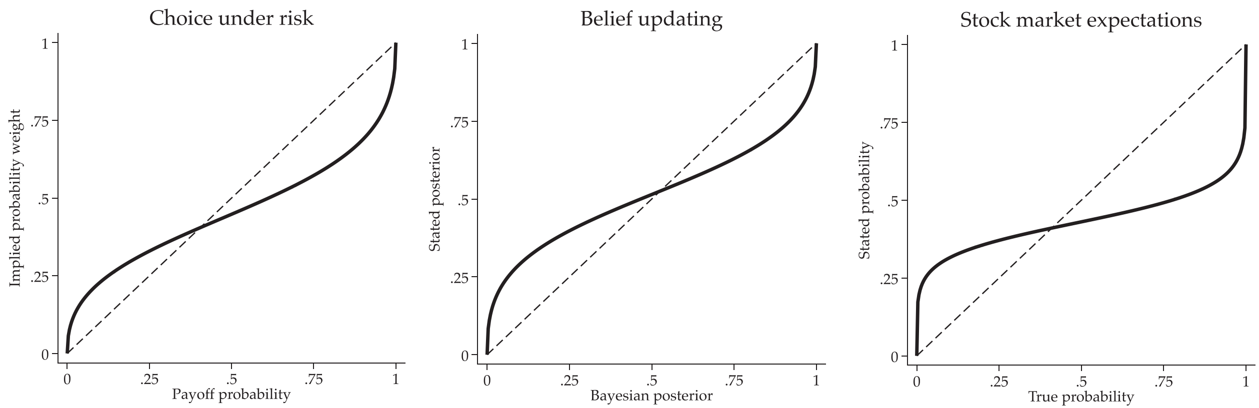

FIGURE I

Decisions as Functions of Objective Probabilities

The left panel illustrates a probability weighting function in choices between monetary gambles. The middle panel illustrates the relationship between stated beliefs and Bayesian posteriors in belief-updating experiments. The right panel shows the typical relationship between stated subjective probabilities and objective (historical) probabilities in surveys about stock returns or inflation. All functions are estimated from the data discussed in Online Appendix E.

alike to some degree, which generates a compression effect to an "intermediate" value.

One view in the literature—reflected in the existence of a large number of dedicated models of probability weighting and belief updating—is that these phenomena reflect domain-specific biases or even preferences. Another view is that they have a common origin: in response to the inherent complexity of forming beliefs and choosing between lotteries, people may use simpler, noisy, or heuristic decision modes. For example, in recent Bayesian cognitive noise models of choice under risk, the difficulty of translating objective probabilities into decisions introduces cognitive noise, which induces the decision maker to partially regress to (or anchor on) an intermediate cognitive default, thus producing probability weighting through a mechanism akin to the classical anchoring-and-adjustment heuristic. Similarly, systematic compression to an intermediate value can result if people choose randomly with some probability. Regardless of the underlying decision mode, this class of models highlights that random noise often generates systematic bias. At the same time, there is little evidence that directly ties together and explains behavior across the three decision domains in Figure I as a function of noisy cognition and complexity.

To make progress, we measure cognitive uncertainty as a proxy for the inherent noisiness or heuristic nature of people's decisions. We conduct a series of online experiments with a total of more than 3,000 participants. We elicit entirely standard controlled decisions in the three domains discussed above. In addition to eliciting payoff-relevant choices and beliefs, we measure cognitive uncertainty. For example, in lottery valuation tasks, after we elicit a participant's certainty equivalent, we ask them how certain they are (in percent) that their true valuation of the lottery actually lies within a $1 window around their stated valuation. Similarly, after participants state probabilistic beliefs in canonical belief-updating experiments, we ask them how certain they are that the Bayesian posterior is contained in a 2 percentage point window around their stated belief. These questions elicit people's subjective percent chance that their decision is actually (close to) the ex ante utility-maximizing one.

This cognitive uncertainty elicitation has five main features. (i) The measure admits a direct theoretical interpretation of awareness of noise. (ii) As documented by our three applications, the elicitation can be tweaked in minor ways to be applicable to a broad set of decision domains with very different experimental paradigms and elicitation protocols. (iii) The question is a composite measure that potentially captures people's awareness of a multitude of cognitive imperfections, such as imperfect perception, preference uncertainty, problems in integrating utils and probabilities, lack of knowledge of Bayes's rule, computational difficulties, or memory imperfections. As a result, a productive interpretation of cognitive uncertainty is that it captures people's subjective difficulty or perceived complexity of a problem. (iv) The measure is very simple, quick, and costless to elicit, making it easy for researchers to add such a question to their own studies. (v) Cognitive uncertainty is strongly correlated with decision variability in repetitions of the same decision problem, which is a key choice signature of (cognitive) noise.

We find large variation in cognitive uncertainty in all of the decision domains. In choice under risk, more than 80% of all decisions are associated with strictly positive cognitive uncertainty, and this number rises to more than 90% in belief updating. Participants appear relatively consistent in their degree of cognitive uncertainty, both across repeated decisions in the same domain ($r \approx 0.7$) and across different decision domains.

Measured cognitive uncertainty strongly predicts observed choices and beliefs in a way that sheds light on the empirical anomalies summarized in Figure I. In all three decision domains, high cognitive uncertainty decisions are substantially more compressed and less responsive to variation in objective probabilities. For example, in choice under risk, high cognitive uncertainty decisions exhibit a substantially shallower slope of the probability weighting function, which implies that cognitive uncertainty is strongly correlated with the well-known fourfold pattern of risk attitudes. For decisions with cognitive uncertainty of zero, the median decision exhibits essentially no probability weighting.

In the domains of beliefs and expectations, we likewise see that high cognitive uncertainty beliefs are substantially more compressed toward 50:50. This means that cognitively uncertain people will sometimes appear more optimistic and sometimes more pessimistic than is warranted, purely depending on whether the true probability is high or low. Cognitive uncertainty is also strongly predictive of more structural belief-updating biases, including base rate insensitivity and conservatism. We discuss implications of these results for interpreting heterogeneity in economic expectations surveys.

The predictive power of cognitive uncertainty for compression effects in decisions is not only driven by the extensive margin of cognitive uncertainty. Instead, the link is strictly monotonic: people in the lowest cognitive uncertainty quartile respond more to objective probabilities than people in the second quartile, who in turn respond more than those in the third quartile, and so on. This shows that the magnitude of cognitive uncertainty contains much information even away from the rational benchmark of zero, and that strictly positive cognitive uncertainty is not just driven by measurement error.

We are agnostic over whether the strong correlations between cognitive uncertainty and behaviors reflect a causal effect of the true (cognitive or decision) noise that underlies cognitive uncertainty or whether awareness of potential errors itself drives behaviors. Under either interpretation, our hypothesis is that the link between cognitive uncertainty and decisions partly reflects the complexity of identifying the utility-maximizing decision. To directly investigate this complexity interpretation, we implement different treatments that vary the complexity of the lottery valuation and belief-updating tasks. In one set of experiments, we vary the computational complexity of the decision problems by displaying the relevant problem parameters (such as payout probabilities or base rates) as algebraic expressions. In other experiments, we increase problem complexity by turning lotteries or belief-updating tasks into compound (multistage probabilistic) problems.

We always find that higher complexity leads to greater cognitive uncertainty, which lends credence to our interpretation that cognitive uncertainty partly reflects the subjectively perceived complexity of decision problems. Moreover, the compression effects summarized in Figure I become substantially more pronounced in the more complex treatments. For instance, contrary to the predictions of (cumulative) prospect theory, the probability weighting function exhibits substantially stronger likelihood insensitivity when the decision problems are more complex. Similarly, in contrast to models of base rate neglect or conservatism that rest on assumptions of fixed parametric biases, the magnitude of base rate insensitivity and conservatism strongly depends on the complexity of the decision problem.

To sum up, this article documents that cognitive uncertainty can be effectively used to test hypotheses about cognitive or decision noise that are difficult to test otherwise. Our results highlight that various judgment and decision errors that are traditionally viewed as distinct share common cognitive origins: the noisy or heuristic decision making people engage in when they find a problem too complex to solve precisely. This insight encourages further research aimed at tying together seemingly distinct behavioral economics anomalies by focusing on the noise that is triggered by complexity. We believe a helpful tool in this regard will be to routinely measure cognitive uncertainty in experiments and surveys, especially given that it is fast and costless to do.

Our work relates to a growing interdisciplinary literature documenting that people often have an awareness of the noisiness of their choices, memories, and perceptions and that they make decisions that are in line with such awareness (e.g., Butler and Loomes 2007; De Martino et al. 2013, 2017; Cubitt, Navarro-Martinez, and Starmer 2015; Drerup, Enke, and Von Gaudecker 2017; Polania, Woodford, and Ruff 2019; Honig, Ma, and Fougnie 2020; Xiang et al. 2021). Our main contribution to this literature is to document that cognitive uncertainty predicts biases across various economic decision tasks and that it can be used to tie together anomalies that are typically viewed as distinct.

This article builds on a broad theoretical literature that has linked probability weighting and over- and underestimation of probabilities to different versions of noise. This includes the recent literature on Bayesian models of cognitive noise (Gabaix and Laibson 2017; Gabaix 2019; Woodford 2020; Frydman and Jin 2022), in particular the model of probability weighting in Khaw, Li, and Woodford (2021).[1] Other noisy decision models of probability weighting and over- and underestimation of probabilities include Viscusi (1985, 1989); Erev, Walsten, and Budescu (1994); Blavatskyy (2007); Bhatia (2014); Marchiori, Di Guida, and Erev 2015; Zhang, Ren, and Maloney (2020); and Erev, Wallsten, and Budescu (1994).[2] Despite the abundance of such models, leading recent reviews rarely even mention a potential role of (cognitive) noise for the empirical regularities and instead emphasize models with fixed "probability weighting," "conservatism," or "extreme belief aversion" parameters that are partly even meant to capture preferences (e.g., Fehr-Duda and Epper 2012; Benjamin 2019).

1. Khaw, Li, and Woodford (2021) and Frydman and Jin (2022) also report experiments on cognitive noise and risk taking, but these do not test predictions related to probability weighting.

2. Wakker (2010) likewise speculates that likelihood insensitivity in probability weighting reflects cognitive limitations. Erev et al. (2017) highlight how an "equal weighting" tendency leads to probability weighting.

O'Donoghue and Somerville (2018) note that "the psychology of probability weighting is poorly understood." This view in the literature may reflect that few contributions directly measure noise or attempt to explain behaviors across different decision domains—both of which we contribute here.[3]

The article proceeds as follows. Section II discusses theoretical background. Section III presents the experimental design. Sections IV–VII discuss the results, and Section VIII concludes.

II. Theoretical Considerations and Hypotheses

Various contributions have hypothesized that the patterns summarized in Figure I are driven by different types of noise. Khaw, Li, and Woodford (2021) model a decision maker who exhibits cognitive noise when processing probabilities, which makes him regress toward an intermediate prior, thus producing probability weighting (also see Gabaix 2019). Earlier related theoretical work modeled probability weighting as resulting from Bayesian updating from imperfect information about objective payout probabilities (Viscusi 1989; Fennell and Baddeley 2012), decision or sampling noise (Blavatskyy 2007; Bhatia 2014), affective versus deliberate decision making (Mukherjee 2010), or random fluctuations in risk preferences (Bhatia and Loomes 2017). Similarly, multiple contributions have argued that regression of beliefs toward 50:50 may reflect noise or ignorance (Viscusi 1985; Erev, Wallsten, and Budescu 1994; Fischhoff and Bruine De Bruin 1999; Marchiori, Di Guida, and Erev 2015; Moore and Healy 2008).

Our analysis builds on these models. We present a stylized adaptation that illustrates how we think about the commonalities reflected in Figure I. Our exposition builds on the recent Bayesian cognitive-noise literature (e.g., Heng, Woodford, and Polania 2020; Woodford 2020; Khaw, Li, and Woodford 2021), though our interpretation of these models is more agnostic.

II.A. Overview

We consider situations in which a decision maker (DM) with Bernoulli utility function $u(\cdot)$ is tasked with making a decision

3. In a paper subsequent to ours, Oprea (2022) provides further evidence that probability weighting is driven by complexity by showing that the fourfold pattern of risk attitudes also holds when risk is removed from lottery choice problems. He reports that these patterns are strongly correlated with cognitive uncertainty.

a that depends on some objective probability p. We denote by $a^*(p) \in \underset{a}{argmax}\, EU(\cdot)$ the DM's true expected-utility-maximizing decision. We assume that through deliberation, the DM only has access to a noisy mental simulation of $a^*(p)$. The noisiness of this mental simulation may depend on the complexity of the decision problem.

II.B. Risky Choice

The DM is asked to indicate his certainty equivalent for a lottery that pays \$1 with probability p and nothing otherwise. By standard arguments, normalizing $u(1) = 1$, the expected-utility-maximizing decision is given by $a^* = u^{-1}(p)$.

II.C. Belief Formation

In a fully controlled "balls-and-urns" belief-updating task, the DM forms beliefs about a binary state of the world, R or B. The DM has prior $b = P(R)$ and receives a binary signal (H or L) with diagnosticity $h = P(H|R) = P(L|B)$. The Bayesian posterior belief is given by $p \equiv P(R|H) = P(B|L) = \frac{bh}{bh+(1-b)(1-h)}$. A widely used formulation that we also leverage is a so-called Grether (1980) decomposition, which generates a linear relationship between the Bayesian posterior odds, the prior odds, and the likelihood ratio: $ln\left(\frac{p}{1-p}\right) = ln\left(\frac{b}{1-b}\right) + ln\left(\frac{h}{1-h}\right)$. We assume the incentive structure is such that it is optimal for the DM to report his true beliefs, such that the utility-maximizing decision is given by $a^* = p$.

II.D. Economic Forecasts

The DM forecasts a future binary state of the world, R or B, that corresponds to a real economic quantity not controlled by the experimenter, such as inflation or stock market growth. In the past, the DM potentially received information about this state of the world, which he processes using Bayes's rule as described, to arrive at posterior p. We again assume the incentive structure is such that it is optimal for the DM to report his true beliefs, $a^* = p$.

II.E. Bayesian Cognitive Noise

As noted already, we assume that the DM does not have access to the utility-maximizing decision $a^*(p)$. This could be due to a variety of reasons. In risky choice, the DM may not know his true utility function, may find it cognitively hard to integrate payoff probabilities and utils, or may have noisy perception. In laboratory belief-updating tasks, the DM may not know Bayes's rule or struggle with implementing it computationally. In economic expectations surveys, the DM may have forgotten financial information that he received in the past, or he may struggle with processing the financial information available to him.

Whatever the underlying cognitive foundations, as we lay out formally in Online Appendix A, we assume that the DM has access to a cognitive signal S that is (scaled) binomially distributed with precision N and satisfies $E[S] = a^*(p)$.[4] This cognitive signal could be interpreted as the outcome of a sequential cognitive sampling or deliberation process as in drift diffusion models. Higher cognitive noise corresponds to a less precise binomial signal. Relatedly, we can think of the level of cognitive noise—and, hence, the precision of the binomial signal—as being determined by the complexity of the decision problem. Indeed, we will provide evidence that higher complexity induces more cognitive noise.

Suppose that the DM holds a beta-distributed prior over $a^*(p)$ and that his decision is given by the Bayesian posterior mean over his utility-maximizing decision.[5] We refer to the mean of the prior, d, as the "cognitive default decision." Given signal realization $s \sim f(s \mid a^*(p))$, a Bayesian DM's decision, a^o, can be represented as a convex combination of the cognitive signal and the prior mean,

4. In contrast to Khaw, Li, and Woodford (2021), our framework features cognitive noise at the level of the utility-maximizing decision, rather than of a problem input parameter. We focus on noise in output space because we want to be transparent that neither we nor our empirical cognitive uncertainty measure take a stance on what the source of cognitive noise is; we believe it is likely that there is more than one source. Online Appendix A.4 discusses how similar predictions to the ones we state below emerge when one compares the behavior of a noiseless and a noisy agent in the framework of Khaw, Li, and Woodford (2021).

5. This assumption has two different interpretations. A first one is that the DM chooses the posterior mean as a heuristic strategy. Indeed, it not entirely clear why a DM who cannot determine a Bayesian posterior or a certainty equivalent should be cognitively capable of best responding to relatively involved incentive structures. A second interpretation is that utility is linear and the DM best responds to the incentives present in the experiment. In our belief-updating experiments, the loss function is quadratic, such that the posterior mean is optimal. In our lottery choice experiments, the implied loss function under risk neutrality is the absolute distance, such that the median is optimal. However, with a binomial distribution, using the mean instead of the median is without much loss because the mean of a beta (a, b) variable is $\frac{a}{a+b}$, the mode is $\frac{a-1}{a+b-2}$ and the median lies between the two.

see Online Appendix A:

$$a^o = \lambda(N) \cdot s + [1 - \lambda(N)] \cdot d, \tag{1}$$

$$E[a^o] = \lambda(N) \cdot a^*(p) + [1 - \lambda(N)] \cdot d. \tag{2}$$

Here, the relative weight placed on the cognitive signal, $\lambda(N)$, increases in the signal's precision N. This decision rule is compatible with an anchoring-and-adjustment heuristic (Tversky and Kahneman 1974), according to which people anchor on some initial reaction, d, and then adjust in the direction of the true utility-maximizing decision after deliberation.

We interpret the prior mean d as the decision the DM would make in the absence of any deliberation. We do not provide a theory of what determines the prior. For our purposes, all that matters is that its mean is sufficiently "intermediate" in nature: for low enough p, $a^*(p) < d$, and for large enough p, $a^*(p) > d$. An intermediate prior implies that people's decisions look like they treat different payout probabilities as more similar than they really are, consistent with the emphasis on "equal weighting" in Erev et al. (2017). Indeed, a large literature argues that people's heuristic (or initial) responses to decision problems are intermediate, such as in research on central tendency effects in judgment and perception (e.g., Hollingworth 1910; Petzschner, Glasauer, and Stephan 2015; Xiang et al. 2021), compromise effects in choice (Simonson and Tversky 1992; Beauchamp et al. 2020), and research that interprets 50:50 responses in economic expectations surveys as a manifestation of "I don't know" (Fischhoff and Bruine De Bruin 1999). The prior distribution could also be partly adapted to which decision "makes sense" on average in a given context.[6]

Note that an alternative interpretation of the DM's decision process that is formally very similar to the Bayesian cognitive noise model in terms of its implications for observable decisions is that of random choice.[7]

6. In their study of inverse S-shaped probability and frequency estimates, Zhang and Maloney (2012) report that a $\frac{1}{N}$ formulation (N being the number of states of the world) captures people's cognitive anchor well.

7. This second possible account of equation (1) is that with probability λ the DM deliberates and plays his resulting cognitive signal s, whereas with probability $(1 - \lambda)$ he plays randomly by drawing from a distribution function with mean d. Under this interpretation, the probability of playing randomly increases in the DM's cognitive noisiness. Note that the only difference between the Bayesian

II.F. Discussion

The linear equation (2) corresponds to the widely studied "neo-additive weighting function" that has attracted attention in the literature on choice under risk. Our stylized framework motivates this functional form by endogenizing its parameters: (i) the intercept increases in noise, and (ii) the slope decreases in noise. A characteristic feature of this decision rule is the "flipping" property implied by Figure I. For instance, in lottery valuation tasks, relative to a noiseless DM, a cognitively noisy DM is less risk-averse for low payout probabilities yet more risk-averse for high payout probabilities.

Equation (2) implies an attenuated but linear mapping between objective probabilities and decisions (when utility is linear). As summarized in Figure I, decisions actually tend to be inverted S-shaped functions of objective probabilities. We explore how cognitive uncertainty relates to this phenomenon in Section VIII and Online Appendix D. To foreshadow this discussion, we find that empirically measured cognitive uncertainty is hump-shaped in objective probabilities, which helps us understand why we typically observe a higher sensitivity of responses to probabilities close to the boundaries than at intermediate levels.

II.G. Predictions

Formal statements of predictions and proofs are relegated to Online Appendix A.

i. Cognitive noise and compression effects.
   a. In risky choice, cognitive noise is correlated with probability weighting: $\exists p^*$ such that, for $p < p^*$, certainty equivalents increase in cognitive noise and for $p > p^*$ they decrease in cognitive noise.
   b. In stated beliefs and economic forecasts, cognitive noise is correlated with overestimation of small and underestimation of large probabilities. In Grether decompositions, cognitive noise is correlated with base rate insensitivity and conservatism.

ii. The distance between the DM's decision and the utility-maximizing decision increases in cognitive noise.

---

cognitive noise and random choice interpretations of equation (2) is whether the DM's average action is attenuated because he regresses to a fixed prior or because he chooses randomly. We embrace both interpretations.

II.H. Empirical Implementation: Cognitive Uncertainty

People's actual level of cognitive noise is conventionally unobservable. To render the predictions testable, we make use of the idea that awareness of cognitive noise generates subjectively perceived uncertainty about what the utility-maximizing decision is. This cognitive uncertainty is measurable. In the context of the framework sketched above, we define it as

$$p_{CU} \equiv P(|\ [a^*|S = s] - a^o\ | > \kappa). \tag{3}$$

Here, $a^*|S = s$ denotes the perceived posterior distribution about the maximizing decision, conditional on having received cognitive signal s. Intuitively, cognitive uncertainty captures the likelihood with which the DM thinks his utility-maximizing decision falls outside a window of arbitrary length κ around the decision he actually chose.[8]

As we show in Online Appendix A, cognitive uncertainty decreases in the precision of the binomial cognitive signal. This allows us to use cognitive uncertainty as a proxy for the magnitude of cognitive noise and, hence, λ. Our argument is not that awareness of cognitive noise necessarily causes the economic behavior of interest (although it may) but that it allows for the measurement of a concept that is difficult to quantify otherwise.

III. Experimental Design

III.A. Overview

As summarized in Table I, we implemented two sets of experiments. The main set of experiments reported here, identified by letter A, was run in early 2022. Earlier experiments (B) were run in 2019. We summarize both sets of experiments here and relegate a detailed exposition of the B experiments to Online Appendix E.

III.B. Decision Tasks

1. *Choice Under Risk.* To estimate a probability weighting function, treatment *Risk A* elicited certainty equivalents for binary lotteries that paid $\$y \in \{15, 16, \ldots, 25\}$ with probability

8. In empirical implementations, κ should be chosen so that the resulting measurement picks up as much variation as possible. This implies that the choice of κ depends on the response scale and should be neither too small nor too large, to avoid bunching at 1 or 0, respectively.

TABLE I
Overview of Experiments

Experiment	Components	# Particip.	Pool
Risk A	Baseline risky-choice tasks (gains) Complex numbers manipulation	500	Prolific
Beliefs A	Baseline belief-updating tasks Complex numbers manipulation	500	Prolific
Risk B	Baseline risky-choice tasks (gains and losses) Compound lottery manipulation	700	AMT
Beliefs B	Baseline belief-updating tasks Compound belief manipulation	700	AMT

Notes. All experiments elicited expectations about the one-year return of the S&P 500, and the B experiments additionally measured expectations about one-year inflation rates and the national income distribution. AMT, Amazon Mechanical Turk.

$p \in \{1, 5, 10, 25, 35, 50, 65, 75, 90, 95, 99\}$ percent, and nothing otherwise. Certainty equivalents were elicited using the BDM technique proposed by Healy (2018). Participants were instructed that for each lottery there is a list of questions that ask whether the participant prefers the lottery or a safe payment, where the safe payment increases as one goes down the list. Following Healy (2018), instead of asking participants to make a decision in every row of the list, we instructed them that they would tell us the safe amount at which they would switch from preferring the lottery to preferring the safe payment and that we would then fill out the entire choice list based on their decision. Thus, participants simply entered a dollar amount into a text box to indicate their certainty equivalent, where entries were restricted to be between zero and the lottery upside. Each participant initially stated their valuation of six randomly selected lotteries.

The two main advantages of this design are that (i) it eliminates the need to go through a long choice list that may be mentally tiring for participants, and (ii) it is well known that the choice list procedure has its own effects on behavior (e.g., Beauchamp et al. 2020), and we wanted to ensure that our results on cognitive uncertainty do not just capture such choice list effects.

In treatment *Risk B*, on the other hand, we instead implemented standard choice lists of the type used by, for example, Tversky and Kahneman (1992); Bruhin, Fehr-Duda, and Epper (2010); and Bernheim and Sprenger (2020). The fact that the results turn out to be very similar suggests that the elicitation technique as such does not generate our results.

We often work with a simple linear transformation of elicited certainty equivalents, normalized certainty equivalents, which are given by the certainty equivalent divided by the upside of the lottery (a quantity that is by construction between 0% and 100%).

2. *Belief Updating.* In designing a structured belief-updating task, we follow the recent review by Benjamin (2019) and implement the workhorse paradigm of so-called balls-and-urns or "bookbags-and-pokerchips" experiments. In treatment *Beliefs A*, there are two bags, A and B. Both bags contain 100 balls, some of which are red and some of which are blue. The computer randomly selects one of the bags according to a prespecified base rate. Subjects do not observe which bag was selected. Instead, the computer selects one or more balls from the selected bag at random (with replacement) and shows them to the subject. The subject is then asked to state a probabilistic guess that either bag was selected. We visualized this procedure for subjects using the image in Online Appendix Figure 8.

The three key parameters of this belief-updating problem are (i) the base rate $b \in \{1, 5, 10, 30, 50, 70, 90, 95, 99\}$ (in percent), which we operationalized as the number of cards out of 100 that had "bag A" as opposed to "bag B" written on them; (ii) the signal diagnosticity $d \in \{65, 75, 90\}$, which is given by the number of red balls in bag A and the number of blue balls in bag B (we only implemented symmetric signal structures such that $P(red|A) = P(blue|B)$); and (iii) the number of randomly drawn balls $M \in \{1, 3, 5\}$. These parameters were randomized across trials but always known to participants.

Each subject initially completed six belief-updating tasks. Financial incentives were implemented through the binarized scoring rule (Hossain and Okui 2013). Here, the probability of receiving a prize of \$10 was given by $\pi = \max\{0, 1 - 0.0001 \cdot (g - t)^2\}$, where g is the guess (in %) and t is the true state (0 or 100).

3. *Economic Forecasts.* All of our experiments also elicited forecasts of economic variables such as stock market returns. A conceptual difference between expectations about real-life quantities and the types of experimental tasks summarized above is that in the latter the experimenter supplies all information that the subject needs to make a well-defined rational decision, while in expectations surveys the experimenter does not have access to

the respondent's information set. Still, cognitive uncertainty can be measured in a similar way. Indeed, intuitively, people may well exhibit cognitive uncertainty about their economic expectations: they may not perfectly remember their current beliefs about the stock market (or the information they received in the past), or they may worry that they have incorrectly processed past information.

In our A study ($N = 1{,}000$, see Table I), we elicit probabilistic forecasts of the performance of the S&P 500. Because incentivizing expectations about future events creates various logistical issues such as credibility concerns and the necessity to wait for future variables to have realized, we elicited them without financial incentives. This is in line with the vast majority of the literature on survey expectations. Each participant responded to the following question:

> The S&P 500 is an American stock market index that includes 500 of the largest companies based in the United States. Jon invested \$100 in the S&P 500 today. What is the percent chance that the value of his investment will be less than \$$y$ in one year from now?

Across participants, the value of y was drawn at random from the set $\{62, 77, 90, 100, 112, 123, 127, 131, 134\}$. These values were chosen such that the corresponding historical return probabilities (from 1980 to 2018) vary between 1% and 99%. For example, the historical probability that a \$100 investment will be worth less than \$127 one year later is 75%. In our "B" experiments, we also elicited beliefs about future inflation rates and the national income distribution in a very similar manner, see Online Appendix E.

III.C. Measuring Cognitive Uncertainty

1. *Elicitation.* In all decision tasks summarized above, decisions are given by a scalar. Loosely speaking, we always measure cognitive uncertainty (CU) on the subsequent screen by eliciting the participant's subjective probability that their expected-utility-maximizing decision is contained in a window around their actual decision.

In choice under risk, we reminded participants of the lottery they were exposed to on the previous screen and then asked:

> Your decision on the previous screen indicates that you value this lottery as much as receiving \$$x$ with certainty. How certain are

> you that you actually value this lottery somewhere between getting $\$(x-0.50)$ and $\$(x+0.50)$?

Participants answered this question by selecting a radio button between 0% and 100%, in steps of 5%. Online Appendix G.1 provides screenshots. In line with the discussion in Section II, we interpret this question as capturing the participant's (posterior) uncertainty about their utility-maximizing decision, after some sampling of cognitive signals has taken place. We refer to inverted responses to this question as cognitive uncertainty rather than confidence because in economics the latter is commonly used for problems that have an objectively correct solution.

In belief updating, the instructions introduced the concept of an "optimal guess." This guess, we explained, uses the laws of probability to compute a statistically correct statement of the probability that either bag was drawn, based on Bayes's rule. We highlighted that this optimal guess does not rely on information that the subject does not have. After indicating their probabilistic belief, subjects were asked (see Online Appendix Figure G.2):

> Your decision on the previous screen indicates that you believe there is an $x\%$ chance that Bag A was selected. How certain are you that the optimal guess is somewhere between $(x-1)\%$ and $(x+1)\%$?

In economic forecasts, the elicitation is very similar, asking how certain the respondent is that their probabilistic guess is within a one percentage point band around the guess that's optimal given the information available to the respondent. Thus, the question does not elicit people's awareness of their lack of information, but instead their perceived ability to appropriately remember and process the information available to them (see Online Appendix Figure G.3):

> On the previous screen, you indicated that you think there is an $x\%$ chance that a \$100 investment into the S&P 500 today will be worth less than $\$y$ in one year from now. How certain are you that the statistically optimal guess (given the information you have) is somewhere between $(x-1)\%$ and $(x+1)\%$?

The biggest difference between our A experiments and the B experiments conducted earlier is the wording of the CU question. In the B experiments, we did not elicit participant's subjective probability that the utility-maximizing decision is within some fixed band around their actual decision, but rather a heuristic confidence interval. In choice under risk, subjects used a slider to calibrate the statement “I am certain that the lottery is worth between a and b to me.” If the participant moved the slider to the very right, a and b corresponded to the previously indicated certainty equivalent. For 20 possible ticks that the slider was moved to the left, a decreased and b increased by 25 cents, in real time. In belief-updating questions and economic forecasts, subjects navigated a slider to calibrate the statement “I am certain that the optimal guess [economic forecasts: statistically optimal guess] is between a and b,” where a and b collapsed to the subject’s own previously indicated guess in case the slider was moved to the very right. For the 30 possible ticks that the slider was moved to the left, a decreased and b increased by 1 percentage point. See Online Appendix E for further details. We believe the new measure to be superior in that it admits a direct quantitative interpretation and is more intuitive for subjects. This being said, the results are qualitatively very similar across both sets of experiments.

2. *Potential Origins of CU.* Our measure is deliberately designed to capture participants’ overall subjective uncertainty about their utility-maximizing decision. This uncertainty could have various potential origins. In choice under risk, people may have imperfect perception, may not know their true preferences, or struggle with integrating utils and probabilities. In belief updating, participants may not know the normatively correct updating rule, or struggle with its computational implementation. In survey expectations, they may not remember information they received in the past, or may again implement an incorrect updating rule. While we conjecture that it will often be of secondary interest to economists what the source of cognitive noise is (there are likely many), we caution that our measure does not allow researchers to directly test models that take a direct stance on the source of the noise.

3. *Comparison with Alternative Measures.* Broadly speaking, the literature has proposed two different types of measures for eliciting people’s uncertainty about their own decisions. At one extreme, psychologists, neuroscientists, and some economists elicit measures of “decision confidence,” in which subjects indicate on Likert scales how confident or certain they are in their decision (e.g., Butler and Loomes 2007; De Martino et al. 2013, 2017; Drerup, Enke, and Von Gaudecker 2017; Polania, Woodford, and Ruff 2019; Xiang et al. 2021). At the other extreme, economists have

used measures of across-trial variability in choices (Khaw, Li, and Woodford 2021) or deliberate randomization (Agranov and Ortoleva 2017, 2020). Our preferred measure strikes a middle ground between these two approaches. Although our approach retains the attractive simplicity of implementing a single question (as in the psychology literature), it also admits a direct quantitative interpretation in terms of a subjective percent chance.[9] The simplicity of asking one question per decision should be contrasted with the approach of gauging cognitive noise through across-task variability in choices, which requires many trials and is often defined at the level of a study rather than of a single choice problem.

4. *Financial Incentives and Validation.* We deliberately do not financially incentivize our elicitation of CU, for two reasons. First, an additional scoring rule makes the measure itself more complex, which increases the cognitive burden on participants. Indeed, recent work documents that unincentivized measures of beliefs are sometimes superior to incentivized ones because they reduce the strategic incentives to game a potentially complex (and misperceived) scoring rule (Danz, Vesterlund, and Wilson 2020). Second, we believe that financially incentivizing the measurement in potentially complicated ways would increase the costs for future researchers to include a CU measure in their experiments and surveys.

We validate our simple but unincentivized measure below by documenting correlations with across-trial variability in repetitions of the same decision problem, which is commonly viewed as a key signature of cognitive noise.

III.D. Complexity Manipulations

Our experiments link cognitive noise to decisions in two ways. First, we correlate decisions with cognitive uncertainty (awareness of noise). Second, we exogenously manipulate the noisiness of decisions by making the decision tasks more complex. In doing so, we focus on the choice under risk and balls-and-urns belief-updating experiments because they allow for more controlled variation.

9. We have found that economists are often more comfortable with uncertainty questions that have a precise quantitative meaning in terms of probabilities, which Likert scales do not.

1. *Complex Numbers.* In our main experiments, *Risk A* and *Beliefs A*, the complexity manipulation is given by representing payout probabilities (in choice under risk) and base rates / signal diagnosticities (in belief updating) as mathematical expressions, such as "Get \$20 with probability $(\frac{7\times6}{2} - 11)\%$." These treatments were implemented in a between-subjects design: after each subject had completed six baseline tasks of either risky choice or belief updating, for a second set of six tasks they were randomized into another set of baseline tasks or a set of the complex numbers tasks.

2. *Compound Problems.* In our experiments *Risk B* and *Beliefs B*, we manipulated complexity by deploying compound problems. We hypothesize that these are more complex for people to think through than the normatively identical reduced problems. The compound problems were randomly interspersed with the respective baseline problems in a within-subjects design. In choice under risk, if a baseline lottery is given by a $p\%$ chance of getting \$20, then the corresponding compound lottery is to get \$20 with probability $p' \sim U\{p - 0.05, \ldots, p + 0.05\}$. In terms of implementation, we told participants that the probability of receiving the lottery upside was unknown to them and would be randomly determined by drawing from a known interval, such that each integer is equally likely to get drawn. Because expected utility is linear in probabilities, this compound manipulation does not affect the normative benchmark for behavior.

In belief updating, if a baseline updating problem features signal diagnosticity h and base rate $b = 50\%$, then the corresponding compound updating problem features diagnosticity $h' \sim U\{h - 0.1, \ldots, h + 0.1\}$. It is straightforward to verify that the Bayesian posterior for these two updating problems is identical.

III.E. Experiments A and B

We summarize the main differences between treatments *Risk A* and *Beliefs A* on the one hand, and *Risk B* and *Beliefs B* on the other hand. (i) The CU measurement differs in wording and quantitative interpretation. (ii) The risky choice tasks were implemented using different procedures: with a BDM mechanism à la Healy (2018) in the A experiments and as a visual multiple price list in the B experiments. (iii) The complexity manipulations differ. Moreover, these were implemented in a between-subjects format in the A experiments and a within-subjects format in the B experiments. (iv) The A experiments feature some repeated, identical problems that allow us to study choice variability. (iv) The B experiments include a broader set of questions measuring economic forecasts.

III.F. Logistics and Participant Pool

As summarized in Table I, our A experiments were conducted on Prolific, while the B experiments were run on Amazon Mechanical Turk (AMT). The B experiments were preregistered, see Online Appendix E.

In both sets of experiments, we took two measures to achieve high data quality. First, our financial incentives are unusually large both by AMT and Prolific standards. Average hourly earnings in our experiments exceed the target compensation on those platforms by roughly 190% and 250%, respectively. Second, we screened out inattentive prospective subjects through comprehension questions and attention checks. In total, 53% and 54% of all prospective participants were screened out in experiments *Risk* and *Beliefs*, respectively. Screenshots of instructions and comprehension check questions can be found in Online Appendix G.

The timeline of *Risk A* and *Beliefs A* was as follows: (i) main incentivized task; (ii) hypothetical economic forecast question; (iii) incentivized Raven matrices test; (iv) demographic questionnaire. Participants received a completion fee of $3 in both treatments. In addition, each participant potentially earned a bonus. With probability 30%, a randomly selected task of part (i) was payoff relevant and with probability 70% part (iii) was paid out. Average earnings in *Risk A* were $8.10 and $4.80 in *Beliefs A*.

IV. Cognitive Uncertainty: Variation and Validation

IV.A. Variation

Figure II shows histograms of task-level CU in the baseline tasks of *Risk A* and *Beliefs A* and for stock market expectations. The magnitude of CU should not be compared across decision domains because the length of the interval with respect to which CU is measured is not comparable. Rather, we show these histograms side by side to illustrate (i) that a large majority of decisions reflect strictly positive CU and (ii) the large heterogeneity in CU. Eighty-three percent of the certainty equivalents in *Risk A*, 93% of beliefs

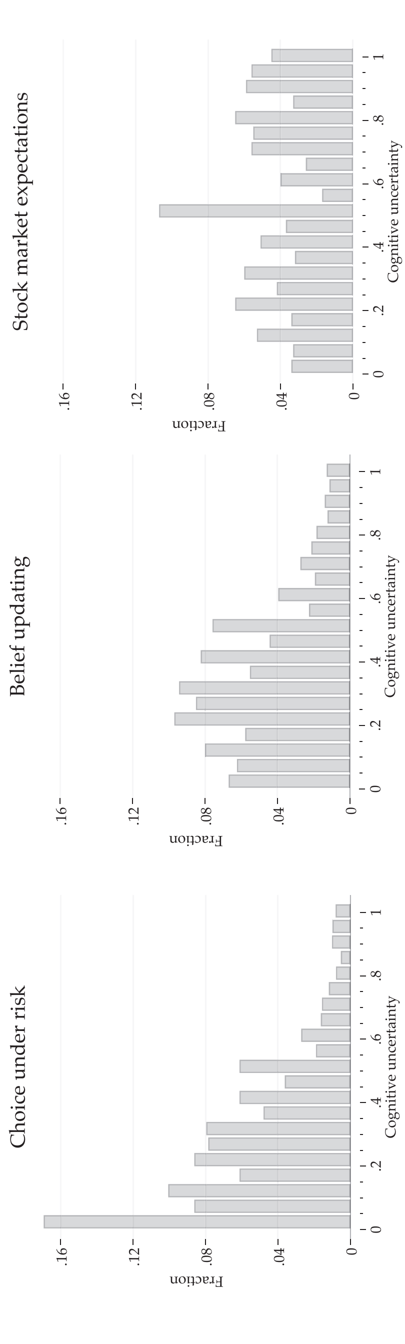

FIGURE II

Histograms of Cognitive Uncertainty in the Baseline Tasks in *Risk A* ($N = 4{,}524$), *Beliefs A* ($N = 4{,}590$), and Stock Market Expectations ($N = 1{,}000$).

in *Beliefs A*, and 97% of stock market forecasts are associated with strictly positive CU.

IV.B. Stability

An obvious question is whether the unincentivized CU question picks up real variation or just noise. A first indication is to look at whether the histograms shown above largely capture within- or across-subject variation. In lottery choice and belief updating, where we observe multiple decisions per subject, 51%–54% of the variation in the CU data is explained by participant fixed effects. Given that some of the residual variation likely reflects measurement error, this suggests that across-subject variation is the dominant source of variation in the CU data, and that participants are relatively consistent in their degree of CU in a given domain.

A second indicator for stability is a within-subject test-retest correlation. This is feasible in our context because in lottery choice and belief updating we implemented at least two decision problems twice. We find that CU is highly correlated across these randomly interspersed elicitations ($r = 0.70$ in *Risk* and $r = 0.68$ in *Beliefs*).

A final indicator for stability is cross-domain stability. We correlate average CU in choice under risk with CU in stock market expectations and average CU in lab beliefs with CU in stock market expectations. The Spearman correlations are given by $\rho = 0.19$ for risky choice and $\rho = 0.35$ for belief updating ($p < .01$ for both correlations). This further suggests some within-person stability of CU.

IV.C. Correlates

Regarding demographic correlates of CU, the most consistent pattern is that—across all three decision domains—women report about 5–11 percentage points higher CU, akin to a large body of evidence on other domains of confidence (see Table II). We also find that older participants report lower CU, although the quantitative magnitude of this relationship is small. Meanwhile, total response time for the survey and proxies for cognitive ability (score on a Raven matrices test and a college degree) are largely unrelated to CU.

Finally, in lottery choice and belief updating, CU strongly decreases in the extremity of the normative benchmark, that is,

TABLE II

CORRELATES OF COGNITIVE UNCERTAINTY

	Dependent variable: Cognitive uncertainty					
	Choice under risk		Belief updating		Stock market exp.	
	(1)	(2)	(3)	(4)	(5)	(6)
1 if female	0.064***	0.063***	0.054***	0.052***	0.11***	0.11***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Age	−0.0017***	−0.0017***	−0.0016***	−0.0016***	−0.0048***	−0.0048***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ln [time taken for study]	−0.018	−0.018	0.011	0.011	0.034	0.034
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Raven matrices score (0–4)	−0.022**	−0.021**	−0.00076	−0.0014	0.0035*	0.0035*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
1 if college degree	0.00059	0.0023	0.023	0.024	−0.029	−0.029
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Extremity of normative decision		−0.19***		−0.31***		0.00026
		(0.03)		(0.03)		(0.00)
Constant	0.47***	0.52***	0.28**	0.37***	0.36**	0.35**
	(0.14)	(0.14)	(0.12)	(0.12)	(0.16)	(0.16)
Observations	4,524	4,524	4,602	4,602	1,000	1,000
R^2	0.04	0.05	0.02	0.06	0.09	0.09

Notes. OLS estimates; robust standard errors (in parentheses) are clustered at the subject level. The extremity of the normative decision is given by the absolute distance of the normative decision to 50%, where the normative decision is assumed to reflect risk neutrality in lottery choice, Bayesian beliefs in belief updating, and historical probabilities in stock market expectations. * $p < .10$, ** $p < .05$, *** $p < .01$.

the absolute distance of the normative benchmark to 50%. In lottery choice, subjects indicate lower CU if the payout probability is far away from 50%, suggesting that, for example, valuing a lottery with payout probability of 95% is easier than valuing a lottery with payout probability 60%. In belief updating, CU reveals that subjects find it easier to state beliefs for problems that have Bayesian posteriors close to zero or one.

IV.D. Cognitive Uncertainty and Choice Variability

Some researchers have used choice variability as an empirical measure of cognitive noise (e.g., Khaw, Li, and Woodford 2021). We examine the empirical correspondence between our CU question and variability for two reasons. First, data on choice variability are useful for understanding whether people's subjective perception of their own noisiness is roughly accurate. Second, a correlation between CU and choice variability may be seen as validation of our quantitative-but-unincentivized question, in the spirit of recent experimental validation studies in the literature (e.g., Falk et al. 2023).

We compute across-trial variability as the absolute difference in decisions across two repetitions of the same problem. We find that decisions that are associated with higher average CU across the two trials are more variable; see Online Appendix Figure 6. The Spearman correlation is $\rho = 0.27$ in choice under risk and $\rho = 0.30$ in belief updating ($p < .01$ in both data sets). These results resonate with those from our work on cognitive uncertainty in intertemporal choice, in which cognitive uncertainty and across-trial variability in responses are likewise significantly correlated (Enke, Graeber, and Oprea 2023).

V. Results: Cognitive Uncertainty Predicts Bias

V.A. Visual Illustration of Compression Effects

We begin by analyzing the data in the baseline tasks.[10] The left panels of Figure III summarize the link between CU and compression effects in the treatment of probabilities. Both sides are constructed following the same logic: by plotting participants'

10. In both *Risk A* and *Beliefs A*, each subject completed six such baseline tasks, after which half the subjects completed six additional baseline tasks, and the remaining half completed the complex math problems. As a result, the data in this section consist of 12 tasks for some subjects and six tasks for others.

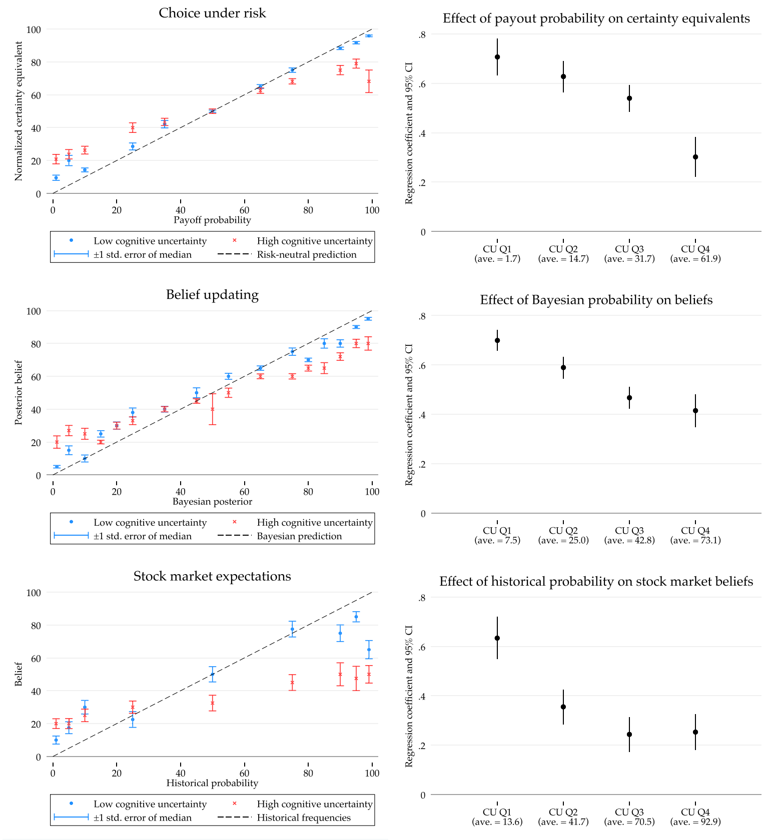

FIGURE III

Cognitive Uncertainty and Sensitivity to Probabilities in Choice under Risk, Belief Updating, and Stock Market Expectations (A Experiments).

Left panels: median normalized certainty equivalents as a function of payout probabilities (top, *Risk A*), median beliefs as a function of binned Bayesian posteriors (middle, *Beliefs A*) and median stock market expectations as a function of historical probabilities (bottom). All panels display bins with 30 or more observations. Low CU is below median. Whiskers show standard error bars. Right panels: coefficients from OLS regressions of (normalized) decisions on objective probabilities, split by CU quartiles. Effect of payout probability on stated certainty equivalents (top, *Risk A*), effect of Bayesian posterior on stated beliefs (middle, *Beliefs A*) and effect of historical probability on stated stock market expectations (bottom). Whiskers show 95% confidence intervals.

(normalized) decisions against objective probabilities. The top row shows normalized certainty equivalents as a function of payout probabilities in *Risk A*. The middle row shows posterior beliefs as a function of Bayesian posteriors in *Beliefs A*. The bottom row shows subjective stock return expectations as a function of historical probabilities. The dots show medians in the samples of above- and below-median CU decisions, respectively.

We see that decisions are always substantially more compressed toward intermediate options in the presence of higher CU. For instance, in choice under risk, the median decision of low CU subjects is frequently visually indistinguishable from the benchmark of no probability weighting. This pattern implies the "flipping" property discussed in the theoretical framework: cognitively uncertain decisions are less risk-averse for low probabilities but more risk-averse for high probabilities. We interpret these patterns as showing that what the literature often refers to as "probability-dependent risk preferences" are, in fact, due to bounded rationality (cognitive noise).

It is instructive to compare the patterns in the top row with those that should be expected from an expected-utility maximizer. As discussed in Section II, normalizing utility from the lottery upside to one, the expected-utility-maximizing decision is given by $a^* = u^{-1}(p)$. Under risk neutrality, normalized certainty equivalents should be located on the 45-degree line. Under strict risk aversion, they should be a convex increasing function of payout probabilities, located strictly below the 45-degree line.

In the belief-updating task, middle row, the median posteriors of low CU decisions are likewise relatively close to the rational benchmark. In contrast, cognitively uncertain beliefs reflect pronounced overestimation of small and underestimation of high probabilities. Thus, the phenomenon of "extreme belief aversion" discussed in the review by Benjamin (2019) reflects cognitive noise rather than preferences.

For the stock market expectations data, bottom row, we plot participants' answers against corresponding historical probabilities. Recall that participants never saw these probabilities—we imputed them from the values of the returns whose probability the participants were asked to assess. Similarly to the lab belief-updating task, we see that cognitive uncertainty is predictive of overestimation of small and underestimation of large probabilities.

The right panels of Figure III provide a more complete picture of the relationship between CU and sensitivity to objective probabilities. We now split the sample into CU quartiles. Because in our lottery choice and belief-updating experiments 20%–25% of all CU statements are equal to zero, the first quartile in these two experiments almost corresponds to $CU = 0$, while the other quartiles leverage variation in the intensive margin of CU. For the four CU buckets, we regress observed (normalized) decisions on the respective objective probability (payout probability in choice under risk, Bayesian posterior in belief updating, and historical probability in stock market expectations) and report the coefficient. If decisions did not depend on cognitive noise, the four regression coefficients would be equally large. Instead, we see that the effect of objective probabilities monotonically decreases as CU increases. This shows that the results are not just driven by the extensive margin of CU but that higher CU is strongly associated with more compression also in the sample of strictly positive CU.

V.B. Regression Evidence

1. *Choice under Risk.* Table III studies the link between CU and likelihood insensitivity (probability weighting) in risky choice more formally, through regression analyses. We estimate the neo-additive weighting function in equation (2). To this effect, we regress certainty equivalents on the payout probability, cognitive uncertainty, and their interaction. The framework in Section II predicts that (i) the interaction coefficient is negative (indicating a shallower slope), and (ii) the raw cognitive uncertainty term is positive, indicating a higher intercept.

Columns (1) and (2) of Table III document that both predictions are indeed borne out in the data. In quantitative terms, an increase in cognitive uncertainty from 0% to 50% is associated with a decrease in the slope of certainty equivalents with respect to payout probabilities by 33.5 percentage points, a very large magnitude.

We likewise find that CU is strongly related to the regression intercept, as predicted by the model. In other words, the positive CU raw term does not mean that the probability weighting function of cognitively uncertain subjects has higher elevation on average—it just means that the elevation at $p = 0$ is higher.

Columns (3)–(6) provide further evidence that these patterns imply the characteristic flipping pattern that we anticipated in

TABLE III

COGNITIVE UNCERTAINTY AND LIKELIHOOD INSENSITIVITY IN *Risk A*

	Dependent variable: Normalized certainty equivalent					
	Full sample		$p < 50\%$		$p \geqslant 50\%$	
	(1)	(2)	(3)	(4)	(5)	(6)
Payout probability	0.73***	0.73***	0.56***	0.55***	0.60***	0.59***
	(0.03)	(0.03)	(0.06)	(0.05)	(0.04)	(0.04)
Payout probability × Cognitive uncertainty	−0.67***	−0.67***				
	(0.08)	(0.08)				
Cognitive uncertainty	25.1***	22.7***	15.0***	11.0**	−26.3***	−27.0***
	(6.18)	(6.02)	(5.64)	(5.48)	(3.51)	(3.60)
Constant	19.7***	31.5***	22.3***	39.4***	30.7***	38.1***
	(2.35)	(4.38)	(2.35)	(6.09)	(3.23)	(4.66)
Demographic controls	No	Yes	No	Yes	No	Yes
Observations	4,524	4,524	2,035	2,035	2,489	2,489
R^2	0.49	0.50	0.10	0.15	0.32	0.32

Notes. OLS estimates; robust standard errors (in parentheses) are clustered at the subject level. Demographic controls include age, gender, college education, and performance on a Raven matrices test. * $p < .10$, ** $p < .05$, *** $p < .01$.

the discussion of the theoretical framework: for small probabilities, cognitively uncertain decisions reflect significantly more risk seeking, while for high probabilities they reflect less risk seeking.

i. *Losses and MPL Elicitation Technique.* Our earlier B experiments allow us to probe the robustness of our results along two dimensions. First, we studied both gain and loss lotteries. Second, the certainty equivalents were elicited using standard visual multiple price lists. The results in these experiments are very similar to those reported already, in the sense that cognitively uncertain decisions are significantly more compressed. This is true for both gains and losses, see Online Appendix E.

The results in the B study imply a nuanced pattern about how CU is correlated with risk-seeking versus risk-averse behavior. Because CU is associated with "overweighting" of small and "underweighting" of large probabilities for gains and losses, we have that high CU decisions reflect risk-seeking behavior for low-probability gains and high-probability losses, but risk-averse behavior for high-probability gains and low-probability losses. In other words, CU is predictive of the so-called fourfold pattern of risk attitudes.

2. *Belief Updating.* Table IV studies the link between CU and belief updating in *Beliefs A*. Again, the framework predicts that cognitive uncertainty should be related to (i) lower sensitivity of beliefs to variation in objective probabilities and (ii) a higher intercept. Columns (1) and (2) directly estimate the neo-additive decision rule in equation (2) that our framework motivates. Here, we link observed beliefs to Bayesian posteriors, CU, and their interaction. Consistent with the visual impression from the left panels of Figure III, cognitively uncertain beliefs are substantially less sensitive to variation in Bayesian posteriors, and their intercept is higher. In terms of quantitative magnitude, the regression coefficients imply that moving from CU of 0% to 50% is associated with a decrease of the slope by 21 percentage points.

i. *Grether Regressions: Inelasticity to Base Rate and Likelihood Ratio (Conservatism).* The literature typically highlights deviations of stated from Bayesian beliefs and the ways people implicitly respond to variation in the base rate, the likelihood ratio, and the sample size (see Benjamin 2019, for a review). As discussed in Section II, we are interested in whether cognitive noise

TABLE IV

COGNITIVE UNCERTAINTY AND BELIEF UPDATING IN *Beliefs A*

	Dependent variable:					
	Posterior belief		ln [Posterior odds]			
	(1)	(2)	(3)	(4)	(5)	(6)
Bayesian posterior	0.71***	0.71***				
	(0.02)	(0.02)				
Bayesian posterior × Cognitive uncertainty	– 0.43***	– 0.43***				
	(0.06)	(0.06)				
Cognitive uncertainty	12.8***	12.7***	– 0.47***	– 0.49***	– 0.48***	– 0.49***
	(3.54)	(3.53)	(0.14)	(0.14)	(0.13)	(0.14)
ln [Bayesian odds]			0.55***	0.55***		
			(0.02)	(0.02)		
ln [Bayesian odds] × Cognitive uncertainty			– 0.42***	– 0.42***		
			(0.07)	(0.07)		
log[Prior odds]					0.69***	0.69***
					(0.03)	(0.03)
log[Likelihood ratio]					0.37***	0.37***
					(0.03)	(0.03)
ln [Prior odds] × Cognitive uncertainty					– 0.52***	– 0.52***
					(0.10)	(0.10)
ln [Likelihood ratio] × Cognitive uncertainty					– 0.21***	– 0.21***
					(0.07)	(0.07)
Constant	19.5***	18.7***	0.23***	0.29**	0.24***	0.28**
	(1.53)	(2.22)	(0.06)	(0.12)	(0.06)	(0.11)
Demographic controls	No	Yes	No	Yes	No	Yes
Observations	4,602	4,602	4,602	4,602	4,602	4,602
R^2	0.49	0.49	0.45	0.45	0.48	0.48

Notes. OLS estimates; robust standard errors (in parentheses) are clustered at the subject level. To avoid a mechanical loss of observations resulting from the log odds definition, the log posterior odds in columns (3)–(6) are computed by replacing stated posterior beliefs of 100% and 0% by 99% and 1%, respectively. The results are virtually identical without this replacement. Demographic controls include age, gender, college education, and performance on a Raven matrices test. * $p < .10$, ** $p < .05$, *** $p < .01$.

could generate the well-known phenomena of base rate insensitivity, conservatism (likelihood ratio insensitivity), and sample size insensitivity.

To analyze this empirically, we resort to Grether regressions (Grether 1980). This specification is derived by expressing Bayes's rule in log form, which implies a linear relationship between the posterior odds, the prior odds, and the likelihood ratio. The canonical finding in the literature is that in these regressions the observed coefficients of the log prior odds and the log likelihood ratio are usually considerably smaller than the Bayesian benchmark of one. As discussed in Section II and shown in Online Appendix A, our stylized cognitive noise model predicts that higher cognitive noise leads to higher insensitivity in these regressions. A simple intuition is that if someone always stated posterior beliefs of 50:50, their sensitivity of beliefs to the base rate and likelihood ratio would be zero.

Table IV, columns (3) and (4) estimate a restricted version of a Grether regression, in which we relate the subject's log posterior odds to the log Bayesian odds. This analysis is instructive because it takes place in log odds space (as motivated by the Grether decomposition), but essentially uses the same variables as in columns (1) and (2). Again, we find that cognitive uncertainty is strongly predictive of the degree of insensitivity of log posterior odds with respect to the Bayesian benchmark.

Finally, columns (5) and (6) estimate a standard Grether regression, except that we also account for interactions with cognitive uncertainty. The negative interaction coefficients show that cognitive uncertainty is strongly related to base rate insensitivity and likelihood insensitivity (conservatism). The quantitative magnitudes of the regression coefficients suggest that, for example, base rate sensitivity decreases from 0.69 with CU of 0% to 0.43 with CU of 50%.[11] These patterns document that (at least a part of) what this literature has identified as base rate neglect, conservatism, and extreme belief aversion are in fact not independent psychological phenomena but instead all generated by cognitive noise and resulting compression effects.

11. The interaction coefficients are larger for the log prior odds than for the log likelihood ratio. We can only speculate about why this is the case. In our experiment, base rates are displayed using sets of cards, while diagnosticities are displayed using urns that are filled with 100 colored balls. We cannot rule out that this difference in the way in which information is presented affects the perceived complexity of these decision parameters or their interaction with cognitive noise.

ii. *Sample Size Effects.* As is well known in the literature, experimental data also reveal systematic variation in stated beliefs conditional on Bayesian posteriors. For instance, for a given base rate, the draw of one blue ball gives rise to the same Bayesian posterior as the draw of two blue balls and one red ball, yet experimental participants consistently update more strongly after observing one blue ball (Benjamin 2019). A common explanation is that subjects update based on sample proportions, while Bayesian updating prescribes updating based on sample differences. Our account of CU also provides an explanation for this pattern. The straightforward reason is that stated CU significantly increases in the sample size, holding the sample difference and the Bayesian posterior fixed (see Online Appendix Table 6). That is, subjects appear to find it easier to form beliefs based on one blue ball than based on two blue balls and one red ball. As a result of this systematic variation in cognitive noise, our account correctly predicts that subjects respond more to the sample difference when the sample size is smaller.

iii. *Earlier Experiments.* All of the patterns summarized above also hold in our earlier B experiments, see Online Appendix E.

3. *Stock Market Expectations.* Online Appendix Table 9 presents regression analyses that confirm the visual impression from Figure III: CU is strongly predictive of the degree to which historical stock returns map into probabilistic forecasts. In our earlier B experiments, we find almost identical patterns for the same measure of stock market expectations. Moreover, we find very similar patterns of cognitive uncertainty predicting compression toward 50:50 also for inflation expectations and beliefs about the income distribution. See Online Appendix E.

V.C. Cognitive Uncertainty and Distance to the Optimal Decision

Thus far, the analyses documented that average decisions are more compressed and further away from normative benchmarks when they are associated with higher cognitive uncertainty. In itself, however, this does not imply that cognitively uncertain decisions are located further away from normative benchmarks, on average. To see this, consider a simple example in which the normatively optimal posterior in a belief-updating task is 80%. Then, the average of stated beliefs of 79% and 77% is located further away from the normative benchmark than the average of beliefs of 60% and 100%, yet the average absolute distance is still smaller in the former case.

Our stylized model predicts that cognitive noise produces stronger compression of the average and that it leads to larger average absolute distances to the normatively optimal decision. We here test this additional prediction. For belief updating, we use the Bayesian posterior as the normative benchmark. For survey expectations, we use historical probabilities. For choice under risk, we assume that subjects' objective is to maximize expected value. However, we have verified that very similar results hold when we infer the "true" utility-maximizing decisions by estimating a population-level constant relative risk aversion (CRRA) parameter.

Figure IV summarizes the results. CU and absolute distances to the normative benchmark are significantly correlated (Spearman's $\rho = 0.31$ in risky choice, $\rho = 0.17$ in beliefs, and $\rho = 0.21$ in stock market expectations, $p < .01$ for all comparisons).

V.D. Measurement Error in Cognitive Uncertainty

A prominent concern regarding the measurement of cognitive or preference constructs in experiments is measurement error (Gillen, Snowberg, and Yariv 2019). In our context, measurement error in the CU elicitation could have two implications. First, CU and certainty equivalents / beliefs could be subject to a form of correlated measurement error that would potentially create a mechanical relationship between the occurrence of strictly positive CU and the sensitivity of decisions to objective probabilities. To illustrate, suppose that all subjects actually exhibit zero cognitive noise. Further suppose that (i) more inattentive subjects are more likely to exhibit random measurement error in the CU elicitation that leads them to state strictly positive CU, and (ii) that this same inattention will also lead subjects to state risky decisions or beliefs that are insensitive to objective probabilities. Under this logic, CU and observed decisions would be mechanically correlated. If this were the case, however, we would expect that CU has no predictive power for decisions in the sample of strictly positive CU. As the right panels of Figure III showed, this is counterfactual as the sensitivity of decisions to delays strongly decreases in CU, even conditional on CU > 0.

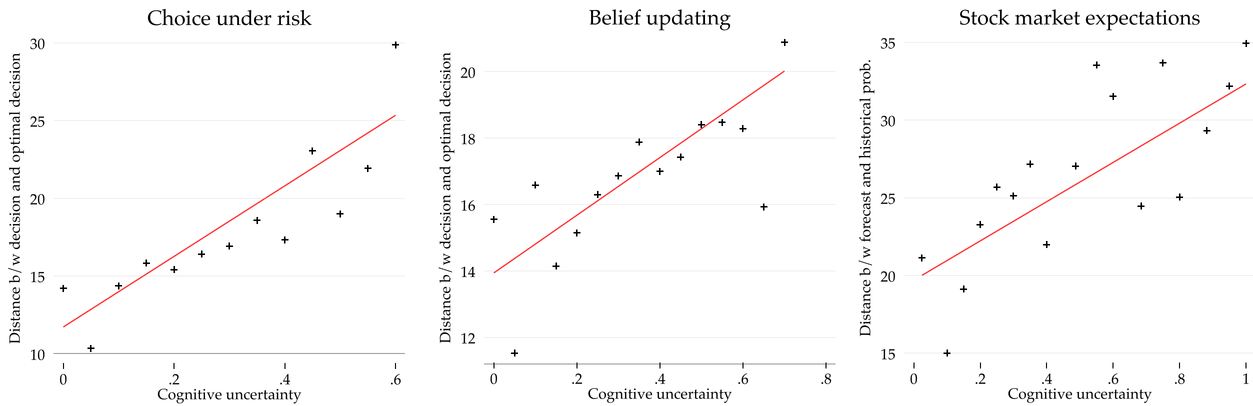

FIGURE IV

Binscatter Plots of Absolute Distance between Decisions and “Normatively Optimal” Decisions as a Function of Cognitive Uncertainty (*A* Experiments)

In the left panel ($N = 4{,}524$) the normative benchmark is assumed to be expected-value maximization, in the middle panel ($N = 4{,}590$) it is the Bayesian posterior, and in the right panel ($N = 1{,}000$) it is historical probabilities. Cognitive uncertainty is winsorized at the 90th percentile in each data set.

A second implication of measurement error in CU could be coefficient attenuation. A standard remedy against this is to instrument out measurement error through repeated elicitations (Gillen, Snowberg, and Yariv 2019). This is feasible in our data because every subject completed at least two decisions twice. As noted already, cognitive uncertainty is highly correlated across these repetitions of the same decision problem ($r = 0.70$ in *Risk A* and $r = 0.68$ in *Beliefs A*). This enables "obviously related instrumental variable" analyses; see Online Appendix Tables 7 and 8. Here, we replicate our OLS regressions from Tables III and IV, except that we instrument for the interaction between objective probabilities and CU with the interaction between objective probabilities and CU from the repeated elicitation. The results are almost identical. This suggests that measurement error in the CU elicitation is not a major concern.

VI. Complexity, Cognitive Noise, and Compression Effects

In the conceptual framework in Section II, we took the magnitude of cognitive noise (captured by N) as given. More realistically, cognitive noise will be higher if the complexity of a decision problem is high. As outlined in Section III, our A experiments manipulated problem complexity by expressing probabilities as math problems. The B experiments instead manipulated complexity through compound problems.

Given that there is no widely accepted theory of what is (not) complex, neither treatment is directly theoretically motivated. However, multiple previous contributions have hypothesized that compound problems or complex numbers can make decision problems harder (e.g., Huck and Weizsäcker 1999; Gillen, Snowberg, and Yariv 2019). Moreover, an added benefit of our CU measurement is that it allows us to directly test whether a complexity intervention actually increases cognitive noise. Both experimental manipulations had large effects on cognitive uncertainty. The complex-numbers manipulation increased CU by 45% in risky choice and by 48% in belief updating. The compound manipulations lead to an increase in CU by 23% in risky choice and by 33% in belief updating.[12]

12. Recall that we used a different CU measure in the B experiments, such that the magnitudes of the CU increase should not be directly compared across experiments.

Figure V documents that this increase in complexity (and resulting cognitive noise) has a large effect on decisions. As predicted, responses are always substantially more compressed toward an intermediate value than in our baseline experiments. This is true for both the math manipulation and the compound problems.[13] Online Appendix Tables 10–13 provide corroborating regression evidence.[14] Overall, we interpret these patterns as evidence that cognitive noise actually causes compression toward an intermediate value, instead of only correlating with it.

We also note that all of these results are inconsistent with a large class of models of probability weighting and belief-updating biases that rest on the assumption of fixed parametric biases, such as base rate neglect parameters or a probability weighting sensitivity parameter. Instead, our results suggest that the complexity of the decision environment partly determines the level of cognitive noise, which in turn drives the magnitude of errors.

VII. Estimating the Central Tendency Effect

The framework laid out in Section II asserts that the compression patterns documented in this paper reflect a regression of average behavior to an "intermediate" d, which could either reflect a fixed default (prior) or the mean random choice. Either interpretation is reminiscent of well-established "central tendency effects" in psychological research on judgment and decision making. Here, we contribute to this discussion by directly estimating the central tendency effect (d), regardless of whether it reflects a fixed prior or the mean random choice. We do not have a general theory of what determines people's priors, though some research in cognitive psychology suggests that the prior may reflect a decision that makes sense on average (e.g., Petzschner, Glasauer, and Stephan 2015; Xiang et al. 2021).[15]

13. It is interesting to relate these results to Harbaugh, Krause, and Vesterlund (2010). They identify evidence for probability weighting in one elicitation mechanism but not another one and interpret this by suggesting that the mechanism that produces probability weighting is "more complex."

14. In experiment *Risk B*, we also implemented compound lotteries for loss gambles. The results are very similar; see Online Appendix E.

15. Some research suggests that people's priors may be influenced by a $\frac{1}{N}$ logic, where N is the number of states (Zhang and Maloney 2012). To test this idea, we ran additional experiments in which we implemented a partition manipulation: in the belief-updating and choice under risk experiments, we increased the number

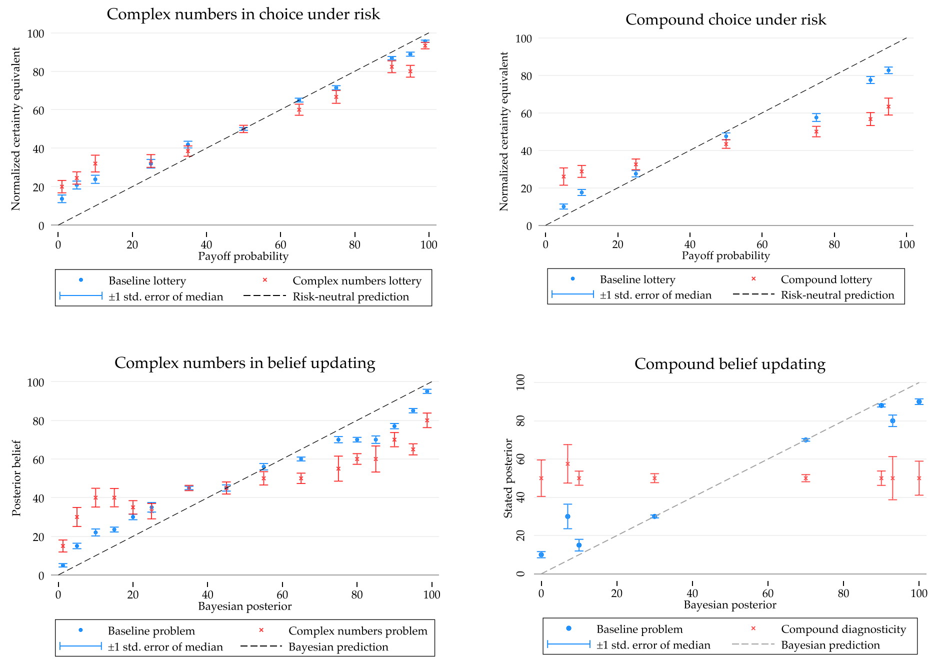

FIGURE V

Complexity and Decisions

Top left panel shows median normalized certainty equivalents separately for baseline and complex-numbers lotteries in the *Risk A* experiment ($N = 3{,}000$). Top right panel shows median normalized certainty equivalents separately for baseline and compound lotteries in the *Risk B* experiment ($N = 1{,}958$). Bottom left shows median posterior beliefs separately for baseline and complex numbers updating problems in the *Beliefs A* experiment ($N = 3{,}000$). Bottom right panel shows median posterior beliefs separately for baseline and compound updating problems in the *Beliefs B* experiment ($N = 2{,}056$). Whiskers show standard error bars. The beliefs figures show bins with more than 30 observations.

Recall that the average decision in our framework can be expressed as a convex combination of the expected-utility-maximizing decision and d, with the relative weight λ being a function of the magnitude of (unobserved) cognitive noise. We proceed by heuristically approximating $\lambda = max\{0; (1 - \gamma p_{CU})\}$, where γ is a nuisance parameter to be estimated. We can then estimate the decision rule in equation (2) as:

$$a^o = \underbrace{max\{1 - \gamma p_{CU}; 0\}}_{\lambda} a^*(p) + \underbrace{min\{\gamma p_{CU}; 1\}}_{1-\lambda} d + \epsilon, \tag{4}$$

where p_{CU} is observed, γ and d are to be estimated, and ϵ is a disturbance term.[16] The utility-maximizing decision a^* is assumed to be the Bayesian posterior in belief updating. For choice under risk, we assume that the utility-maximizing decision reflects CRRA utility, with utility curvature to be estimated.[17]

We estimate this equation at the population level using standard nonlinear least squares techniques. This means that we leverage individual-level (in fact, decision-level) variation in CU but estimate a single average d for the population. For benchmarking purposes, we also estimate a "restricted model" that excludes cognitive noise, that is, setting $p_{CU} = 0$.

Table V reports the model estimates for both our A experiments and the earlier B experiments. There are three main takeaways. First, we consistently estimate an "intermediate" mean of the prior distribution. The estimated cognitive default is very close to 0.5 in the beliefs experiments and somewhat lower at around 0.4 in choice under risk. The estimates of the default decision correspond well with the visual impressions from Figure III and with the large body of work on central tendency or compromise effects in psychology and economics.

The second main takeaway from the model estimations is that allowing for a role of cognitive noise increases model fit

of states from 2 to 10 without changing the normatively relevant features of the problem. Under the assumptions that (i) the model parameter d reflects a fixed prior and (ii) that it is partly influenced by a $\frac{1}{N}$ logic, such a treatment should decrease observed decisions, and more so for cognitively uncertain people. Online Appendix F reports the results of these experiments, which are mixed.

16. Note that in this approach, λ (and hence unobserved cognitive noise) varies at the choice level, but the nuisance parameter γ is fixed at the population level.

17. The estimating equation with CRRA utility curvature parameter α is given by $a^o = max\{1 - \gamma p_{CU}; 0\} p^{\frac{1}{\alpha}} + min\{\gamma p_{CU}; 1\} d + \epsilon$.

TABLE V

ESTIMATES OF CENTRAL TENDENCY EFFECT ACROSS EXPERIMENTS

	Risk A		*Beliefs A*		*Risk B*		*Beliefs B*	
	Restr. (1)	CU (2)	Restr. (3)	CU (4)	Restr. (5)	CU (6)	Restr. (7)	CU (8)
$\hat{d}$	NA	0.43	NA	0.52	NA	0.40	NA	0.52
AIC	18,958	18,477	211	−936	7,996	7,707	211	−935

Notes. Estimates of different versions of equation (4). Columns (1), (3), (5), and (7): set $\gamma = 1$ and $p_{CU} = 0$. All estimated standard errors (computed based on clustering at the subject level) are smaller than 0.02. AIC, Akaike information criterion.

substantially relative to the restricted model that does not include cognitive uncertainty. This can be inferred from the lower values of Akaike's information criterion.

VIII. DISCUSSION

This article has argued that measuring cognitive uncertainty in a simple, fast, and costless manner allows experimental and survey researchers to predict behavior and biases and shed light on the decision modes that underlie commonalities in errors across different domains. Instead of recapitulating our results, we discuss extensions, limitations, and directions for future research.

VIII.A. Extension: S-Shaped Response Functions

While our main empirical analyses focus on the observation that people's beliefs and choices are compressed toward some intermediate value, it is well known in the literature that decisions are often nonlinear (inverse S-shaped) in objective probabilities (see Figure I). As we discuss in detail in Online Appendix D, our account of cognitive uncertainty also sheds light on this regularity. The reason is that CU is hump-shaped in objective probabilities. For example, it appears to be easier for people to value a lottery that has a payout probability close to the boundaries. Similarly, people report lower CU in belief-updating problems that have Bayesian posteriors close to the boundaries. The model estimations in Online Appendix D show that these nonlinearities in how CU depends on objective probabilities can translate into the canonical S-shaped response functions commonly observed in the literature.

VIII.B. Extension: Ambiguity Attitudes

While here we focus on how CU sheds light on the pattern that people treat different objective probabilities alike to some degree, there is also a direct connection to research on ambiguity. The reason is that recent reviews highlight the concept of ambiguity insensitivity, which asserts that people are excessively insensitive to changes in the likelihood of ambiguous events (Trautmann and van de Kuilen 2015). In the working paper version of this paper, we document that measured cognitive uncertainty also strongly predicts the magnitude of ambiguity insensitivity (Enke and Graeber 2019). Indeed, we find that cognitively uncertain people often act as though they are ambiguity seeking when an ambiguous event is very unlikely.

VIII.C. Implications for Research Linking Expectations Measures to Field Behaviors

If stated expectations are systematically distorted due to the types of compression effects that we document in this article, demographic differences in expectations could just reflect heterogeneity in cognitive noise rather than true beliefs. Moreover, when researchers estimate links between expectations and field behaviors, cognitive noise could attenuate these relationships. In line with this conjecture, Drerup, Enke, and Von Gaudecker (2017), Giglio et al. (2019), and Yang (2023) find that the relationship between expectations and investment behavior is considerably more pronounced among people with high confidence in their expectations. We conjecture that CU will be predictive of the strength of the relationship between behaviors and expectations more generally (see also Charles, Frydman, and Kilic 2022; Yang 2023). Thus, at a minimum, measuring CU in surveys allows researchers to conduct heterogeneity analyses regarding the predictability of field behaviors.

VIII.D. Limitations

An obvious limitation of our approach is that we do not have a general theory of what the prior / cognitive default / mean random choice is. We work with the idea that the mean prior reflects the decision they would have made before deliberating about the problem at hand. Yet casual introspection suggests that other factors might also shape people's initial reactions. For instance, if a choice option is displayed in red font, it might be visually salient and therefore serve as a cognitive anchor from which people's deliberation process adjusts.

Related to this discussion is research on bounded rationality that focuses on the role of misleading intuitions, as they result from salience, focusing, or memory-based cueing effects (e.g., Kahneman 2011; Bordalo, Gennaioli, and Shleifer 2013, 2020; Kőszegi and Szeidl 2013; Enke, Schwerter, and Zimmermann 2020). Although this article is more concerned with the effects of complexity than with those of strong intuitions, we conjecture that the (unspecified) cognitive default provides a potential link between these two literatures. We speculate that strong intuitions, salient choice options, or associations-based recall shape people's initial reaction to a choice problem (the prior / cognitive default), while CU captures the degree to which people adjust away from these initial reactions. If true, such a perspective would suggest the testable prediction that salience, focusing, and memory-based cuing effects are particularly pronounced among people with high CU.

More closely integrating cognitive noise with attention and memory research is also relevant because prior work has shown that probability weighting in risky choice and probability estimates are influenced by salience and asymmetric recall (e.g., Stewart, Chater, and Brown 2006; Bordalo, Gennaioli, and Shleifer 2012; Bordalo et al. 2023). Similarly, a broad body of work often identifies the opposite of probability weighting when people decide based on experience rather than problem descriptions (Hertwig and Erev 2009). It is not obvious that our approach of measuring CU can reconcile these patterns.

A third limitation of our work is that we do not have a theory of which aspects of a decision actually generate cognitive noise and resulting CU. As we saw, more complex decisions lead to higher CU. Prior work has shown that cognitive noise is also a function of time pressure, experience, and prior beliefs (Polania, Woodford, and Ruff 2019; Prat-Carrabin and Woodford 2021; Frydman and Jin 2022). Yet a general theory of what makes a task (not) complex is not available. Other aspects that generate cognitive uncertainty may pertain to the decision maker: the availability of cognitive resources or the amount of experience. Future research could helpfully shed light on this.

VIII.E. Open Questions and Potential Applications

We conjecture that the measurement of CU could shed light on behavior in multiple other domains of economic decision making. Most fundamentally, people likely don't just have CU in choosing between lotteries or in updating their beliefs, but also in other domains. For instance, in Enke, Graeber, and Oprea (2023), we study how cognitive uncertainty helps shed light on "anomalies" in intertemporal choice. Yet we speculate that there may be many more applications in which a measurement of CU could shed light on biases and anomalies that have a compression flavor. For example, in the widely studied news vendor game that is of relevance to researchers in economics, management, and operations research, people generally succumb to a pull-to-the-center bias (Schweitzer and Cachon 2000). Similarly, laboratory experiments on effort choice often find that the elasticity of labor supply with respect to piece rates is very low; we again speculate that this insensitivity / compression effect could be explained by measuring CU.

Another open question relates to the link between objective cognitive noise and CU. In the decision contexts that we study here, people's awareness of their own cognitive noise is at least partly accurate. Yet in other decision domains, people's metacognition may be less well calibrated, as in Enke, Graeber, and Oprea (2022). This immediately raises the question of when people's CU is (not) reflective of actual noise.

Finally, another open question concerns the choice implications of cognitive noise. We highlighted the empirical regularity that CU is associated with an attenuated relationship between decisions and problem parameters. In other contexts, CU may predict a form of "caution" (Cerreia-Vioglio, Dillenberger, and Ortoleva 2015) or "complexity aversion," according to which people shy away from choice options regarding which they have high CU. Future research could helpfully shed light on when compression effects or caution dominate.

Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

Data Availability

The data underlying this article are available in the Harvard Dataverse, https://doi.org/10.7910/DVN/IQ47ZB (Enke and Graeber 2023).

Harvard University, United States
Harvard University, United States

References

Agranov, Marina, and Pietro Ortoleva, "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, 125 (2017), 40–68. https://doi.org/10.1086/689774

———, "Ranges of Preferences and Randomization," Princeton University Working Paper, 2020.

Beauchamp, Jonathan P., Daniel J. Benjamin, David I. Laibson, and Christopher F. Chabris, "Measuring and Controlling for the Compromise Effect When Estimating Risk Preference Parameters," *Experimental Economics*, 23 (2020), 1069–1099. https://doi.org/10.1007/s10683-019-09640-z

Benjamin, Daniel J., "Errors in Probabilistic Reasoning and Judgment Biases," in *Handbook of Behavioral Economics: Applications and Foundations*, vol. 2, B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds. (Amsterdam: Elsevier, 2019), 69–186. https://doi.org/10.1016/bs.hesbe.2018.11.002

Bernheim, B. Douglas, and Charles Sprenger, "On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting," *Econometrica*, 88 (2020), 1363–1409. https://doi.org/10.3982/ECTA16646

Bhatia, Sudeep, "Sequential Sampling and Paradoxes of Risky Choice," *Psychonomic Bulletin & Review*, 21 (2014), 1095–1111. https://doi.org/10.3758/s13423-014-0650-1

Bhatia, Sudeep, and Graham Loomes, "Noisy Preferences in Risky Choice: A Cautionary Note," *Psychological Review*, 124 (2017), 678–687. https://doi.org/10.1037/rev0000073

Blavatskyy, Pavlo R., "Stochastic Expected Utility Theory," *Journal of Risk and Uncertainty*, 34 (2007), 259–286. https://doi.org/10.1007/s11166-007-9009-6

Bordalo, Pedro, John J. Conlon, Nicola Gennaioli, Spencer Y. Kwon, and Andrei Shleifer, "Memory and Probability," *Quarterly Journal of Economics*, 138 (2023), 265–311. https://doi.org/10.1093/qje/qjac031

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice under Risk," *Quarterly Journal of Economics*, 127 (2012), 1243–1285. https://doi.org/10.1093/qje/qjs018

———, "Salience and Consumer Choice," *Journal of Political Economy*, 121 (2013), 803–843. https://doi.org/10.1086/673885

———, "Memory, Attention, and Choice," *Quarterly Journal of Economics*, 135 (2020), 1399–1442. https://doi.org/10.1093/qje/qjaa007

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper, "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 78 (2010), 1375–1412. https://doi.org/10.3982/ECTA7139

Butler, David J., and Graham C. Loomes, "Imprecision as an Account of the Preference Reversal Phenomenon," *American Economic Review*, 97 (2007), 277–297. https://doi.org/10.1257/aer.97.1.277

Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva, "Cautious Expected Utility and the Certainty Effect," *Econometrica*, 83 (2015), 693–728. https://doi.org/10.3982/ECTA11733

Charles, Constantin, Cary Frydman, and Mete Kilic, "Insensitive Investors," CESifo Working Paper no. 10067, 2022. https://doi.org/10.2139/ssrn.4273441

Cubitt, Robin P., Daniel Navarro-Martinez, and Chris Starmer, "On Preference Imprecision," *Journal of Risk and Uncertainty*, 50 (2015), 1–34. https://doi.org/10.1007/s11166-015-9207-6

Danz, David, Lise Vesterlund, and Alistair J. Wilson, "Belief Elicitation: Limiting Truth Telling with Information on Incentives," NBER Working Paper no. 27327, 2020. https://doi.org/10.3386/w27327

De Martino, Benedetto, Sebastian Bobadilla-Suarez, Takao Nouguchi, Tali Sharot, and Bradley C. Love, "Social Information Is Integrated into Value and Confidence Judgments According to its Reliability," *Journal of Neuroscience*, 37 (2017), 6066–6074. https://doi.org/10.1523/JNEUROSCI.3880-16.2017

De Martino, Benedetto, Stephen M. Fleming, Neil Garrett, and Raymond J. Dolan, "Confidence in Value-Based Choice," *Nature Neuroscience*, 16 (2013), 105–110. https://doi.org/10.1038/nn.3279

Drerup, Tilman, Benjamin Enke, and Hans-Martin Von Gaudecker, "The Precision of Subjective Data and the Explanatory Power of Economic Models," *Journal of Econometrics*, 200 (2017), 378–389. https://doi.org/10.1016/j.jeconom.2017.06.017

Enke, Benjamin, and Thomas Graeber, "Cognitive Uncertainty," NBER Working Paper no. 26518, 2019. https://doi.org/10.3386/w26518

———, "Replication Data for: 'Cognitive Uncertainty'," (2023), Harvard Dataverse, https://doi.org/10.7910/DVN/IQ47ZB.

Enke, Benjamin, Thomas Graeber, and Ryan Oprea, "Confidence, Self-Selection and Bias in the Aggregate," NBER Working Paper no. 30262, 2022. https://doi.org/10.3386/w30262

———, "Complexity and Time," NBER Working Paper no. 31047, 2023. https://doi.org/10.3386/w31047

Enke, Benjamin, Frederik Schwerter, and Florian Zimmermann, "Associative Memory and Belief Formation," NBER Working Paper no. 26664, 2020. https://doi.org/10.3386/w26664

Erev, Ido, Eyal Ert, Ori Plonsky, Doron Cohen, and Oded Cohen, "From Anomalies to Forecasts: Toward a Descriptive Model of Decisions under Risk, under Ambiguity, and from Experience," *Psychological Review*, 124 (2017), 369–409. https://doi.org/10.1037/rev0000062

Erev, Ido, Thomas S. Wallsten, and David V. Budescu, "Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes," *Psychological Review*, 101 (1994), 519. https://doi.org/10.1037/0033-295X.101.3.519

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *Management Science*, 69 (2023), 1935–1950. https://doi.org/10.1287/mnsc.2022.4455

Fehr-Duda, Helga, and Thomas Epper, "Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences," *Annual Review of Economics*, 4 (2012), 567–593. https://doi.org/10.1146/annurev-economics-080511-110950

Fennell, John, and Roland Baddeley, "Uncertainty Plus Prior Equals Rational Bias: An Intuitive Bayesian Probability Weighting Function," *Psychological Review*, 119 (2012), 878–887. https://doi.org/10.1037/a0029346

Fischhoff, Baruch, and Wändi Bruine De Bruin, "Fifty–Fifty = 50%?" *Journal of Behavioral Decision Making*, 12 (1999), 149–163. https://doi.org/10.1002/(SICI)1099-0771(199906)12:2%3C149::AID-BDM314%3E3.0.CO;2-J

Frydman, Cary, and Lawrence J. Jin, "Efficient Coding and Risky Choice," *Quarterly Journal of Economics*, 137 (2022), 161–213. https://doi.org/10.1093/qje/qjab031.

Gabaix, Xavier, "Behavioral Inattention," in *Handbook of Behavioral Economics: Foundations and Applications*, vol. 2, Douglas Bernheim, Stefano DellaVigna, and David Laibson, eds. (Amsterdam: Elsevier, 2019), 261–343. https://doi.org/10.1016/bs.hesbe.2018.11.001

Gabaix, Xavier, and David Laibson, “Myopia and Discounting,” NBER Working Paper no. 23254, 2017. https://doi.org/10.3386/w23254

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, “Five Facts about Beliefs and Portfolios,” NBER Working Paper no. 25744, 2019. https://doi.org/10.3386/w25744

Gillen, Ben, Erik Snowberg, and Leeat Yariv, “Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study,” *Journal of Political Economy*, 127 (2019), 1826–1863. https://doi.org/10.1086/701681

Grether, David M., “Bayes Rule as a Descriptive Model: The Representativeness Heuristic,” *Quarterly Journal of Economics*, 95 (1980), 537–557. https://doi.org/10.2307/1885092

Harbaugh, William T., Kate Krause, and Lise Vesterlund, “The Fourfold Pattern of Risk Attitudes in Choice and Pricing Tasks,” *Economic Journal*, 120 (2010), 595–611. https://doi.org/10.1111/j.1468-0297.2009.02312.x

Healy, Paul J., “Explaining the BDM—Or Any Random Binary Choice Elicitation Mechanism—To Subjects,” Ohio State University Working Paper, 2018.

Heng, Joseph A., Michael Woodford, and Rafael Polania, “Efficient Sampling and Noisy Decisions,” *eLife*, 9 (2020), e54962. https://doi.org/10.7554/eLife.54962

Hertwig, Ralph, and Ido Erev, “The Description–Experience Gap in Risky Choice,” *Trends in Cognitive Sciences*, 13 (2009), 517–523. https://doi.org/10.1016/j.tics.2009.09.004

Hollingworth, Harry Levi, “The Central Tendency of Judgment,” *Journal of Philosophy, Psychology and Scientific Methods*, 7 (1910), 461–469. https://doi.org/10.2307/2012819

Honig, Maija, Wei Ji Ma, and Daryl Fougnie, “Humans Incorporate Trial-to-Trial Working Memory Uncertainty into Rewarded Decisions,” *Proceedings of the National Academy of Sciences*, 117 (2020), 8391–8397. https://doi.org/10.1073/pnas.1918143117

Hossain, Tanjim, and Ryo Okui, “The Binarized Scoring Rule,” *Review of Economic Studies*, 80 (2013), 984–1001. https://doi.org/10.1093/restud/rdt006

Huck, Steffen, and Georg Weizsäcker, “Risk, Complexity, and Deviations from Expected-Value Maximization: Results of a Lottery Choice Experiment,” *Journal of Economic Psychology*, 20 (1999), 699–715. https://doi.org/10.1016/S0167-4870(99)00031-8

Kahneman, Daniel, *Thinking, Fast and Slow* (New York: Macmillan, 2011).

Khaw, Mel Win, Ziang Li, and Michael Woodford, “Cognitive Imprecision and Small-Stakes Risk Aversion,” *Review of Economic Studies*, 88 (2021), 1979–2013. https://doi.org/10.1093/restud/rdaa044

Kőszegi, Botond, and Adam Szeidl, “A Model of Focusing in Economic Choice,” *Quarterly Journal of Economics*, 128 (2013), 53–104. https://doi.org/10.1093/qje/qjs049

Marchiori, Davide, Sibilla Di Guida, and Ido Erev, “Noisy Retrieval Models of Over- and Undersensitivity to Rare Events,” *Decision*, 2 (2015), 82–106. https://doi.org/10.1037/dec0000023

Moore, Don A., and Paul J. Healy, “The Trouble with Overconfidence,” *Psychological Review*, 115 (2008), 502–517. https://doi.org/10.1037/0033-295X.115.2.502

Mukherjee, Kanchan, “A Dual System Model of Preferences under Risk,” *Psychological Review*, 117 (2010), 243–255. https://doi.org/10.1037/a0017884

O’Donoghue, Ted, and Jason Somerville, “Modeling Risk Aversion in Economics,” *Journal of Economic Perspectives*, 32 (2018), 91–114. https://doi.org/10.1257/jep.32.2.91

Oprea, Ryan, “Simplicity Equivalents,” UC Santa Barbara Working Paper, 2022.

Petzschner, Frederike H., Stefan Glasauer, and Klaas E. Stephan, “A Bayesian Perspective on Magnitude Estimation,” *Trends in Cognitive Sciences*, 19 (2015), 285–293. https://doi.org/10.1016/j.tics.2015.03.002

Polania, Rafael, Michael Woodford, and Christian C. Ruff, “Efficient Coding of Subjective Value,” *Nature Neuroscience*, 22 (2019), 134–142. https://doi.org/10.1038/s41593-018-0292-0

Prat-Carrabin, Arthur, and Michael Woodford, "Bias and Variance of the Bayesian-Mean Decoder," *Advances in Neural Information Processing Systems*, 34 (2021), 23793–23805.
Schweitzer, Maurice E., and Gérard P. Cachon, "Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence," *Management Science*, 46 (2000), 404–420. https://doi.org/10.1287/mnsc.46.3.404.12070
Simonson, Itamar, and Amos Tversky, "Choice in Context: Tradeoff Contrast and Extremeness Aversion," *Journal of Marketing Research*, 29 (1992), 281–295. https://doi.org/10.1177/002224379202900301
Stewart, Neil, Nick Chater, and Gordon D. A. Brown, "Decision by Sampling," *Cognitive Psychology*, 53 (2006), 1–26. https://doi.org/10.1016/j.cogpsych.2005.10.003
Trautmann, Stefan T., and Gijs van de Kuilen, "Ambiguity Attitudes," in *The Wiley Blackwell Handbook of Judgment and Decision Making*, vol. 1, Gideon Keren and George Wu, eds. (Chichester: Wiley, 2015), 89–116. https://doi.org/10.1002/9781118468333.ch3
Tversky, Amos, and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185 (1974), 1124–1131. https://doi.org/10.1126/science.185.4157.1124
———, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (1992), 297–323. https://doi.org/10.1007/BF00122574
Viscusi, W. Kip, "A Bayesian Perspective on Biases in Risk Perception," *Economics Letters*, 17 (1985), 59–62. https://doi.org/10.1016/0165-1765(85)90127-2
———, "Prospective Reference Theory: Toward an Explanation of the Paradoxes," *Journal of Risk and Uncertainty*, 2 (1989), 235–263. https://doi.org/10.1007/BF00209389
Wakker, Peter P., *Prospect Theory: For Risk and Ambiguity* (Cambridge: Cambridge University Press, 2010).
Woodford, Michael, "Modeling Imprecision in Perception, Valuation, and Choice," *Annual Review of Economics*, 12 (2020), 579–601. https://doi.org/10.1146/annurev-economics-102819-040518
Xiang, Yang, Thomas Graeber, Benjamin Enke, and Samuel J. Gershman, "Confidence and Central Tendency in Perceptual Judgment," *Attention, Perception and Psychophysics*, 83 (2021), 3024–3034. https://doi.org/10.3758/s13414-021-02300-6
Yang, Jeffrey, "On the Decision-Relevance of Subjective Beliefs," SSRN Working Paper, 2023. https://doi.org/10.2139/ssrn.4425080
Zhang, Hang, and Laurence T. Maloney, "Ubiquitous Log Odds: A Common Representation of Probability and Frequency Distortion in Perception, Action, and Cognition," *Frontiers in Neuroscience*, 6 (2012), 1.
Zhang, Hang, Xiangjuan Ren, and Laurence T. Maloney, "The Bounded Rationality of Probability Distortion," *Proceedings of the National Academy of Sciences*, 117 (2020), 22024–22034. https://doi.org/10.1073/pnas.1922401117