Cognitive Imprecision and Stake-Dependent Risk Attitudes[*]

Mel Win Khaw
Microsoft Corporation

Ziang Li
Princeton University

Michael Woodford
Columbia University

October 8, 2023

Abstract

In an experiment that elicits subjects' willingness to pay (WTP) for the outcome of a lottery, we document a systematic effect of stake sizes on the magnitude and sign of the relative risk premium, and find that there is a log-linear relationship between the monetary payoff of the lottery and WTP, conditional on the probability of the payoff and its sign. We account quantitatively for this relationship, and the way in which it varies with both the probability and sign of the lottery payoff, in a model in which all departures from risk-neutral bidding are attributed to an optimal adaptation of bidding behavior to the presence of cognitive noise. Moreover, the cognitive noise required by our hypothesis is consistent with patterns of bias and variability in judgments about numerical magnitudes and probabilities that have been observed in other contexts. We thus provide foundations for the kind of nonlinear distortions in lottery valuation posited by prospect theory, that we believe can provide an interpretation for the observed instability across contexts of estimated prospect-theoretic parameters.

[*]Listed affiliations are current ones; Dr. Khaw's work on this project took place while he was a postdoctoral research affiliate of Columbia University. We thank Rahul Bhui, Helga Fehr-Duda, Cary Frydman, Lawrence Jin, Eric Johnson, Ryan Oprea, Charlie Sprenger, Peter Wakker and George Wu for helpful comments, and the National Science Foundation for research support.

One of the more puzzling features of decision making under risk in the laboratory is the fact that the same experimental subjects can display either risk-averse or risk-seeking behavior, depending on the nature of the choices presented to them (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Descriptive models of choice under risk, such as prospect theory, are often understood (especially in the economics literature) as specifying preferences over lottery outcomes, simply preferences that do not conform to the axioms of expected-utility theory. But one cannot explain the differences in apparent risk attitudes across situations — either the fourfold pattern of risk attitudes documented by Tversky and Kahneman (1992) or the alternative fourfold pattern proposed by Markowitz (1952) — except on the hypothesis that what people value is the probability distribution over prospective *changes* in their wealth from a given gamble, rather than the probability distribution over final wealth levels if they take the gamble.[1] Yet it remains unclear why evolution should have given us a brain that assigns value to wealth changes as opposed to the state of having a certain amount of wealth; such valuations are inconsistent with the principle that the value attributed to receiving money should be derived from the value of what one can buy with it.

A recent literature proposes instead that apparent departures from risk-neutrality, at least in laboratory experiments involving stakes that are small relative to a subject's overall budget, actually reflect an efficient adaptation of subjects' decision rules to the presence of cognitive noise.[2] In analyses of this kind, decision rules are posited to be optimal in the sense of maximizing the expected financial wealth of the decision maker (DM).[3] Under this assumption about the ultimate objective served by the decision rules that people have learned to use, there is no difference between selecting rules on the basis of the mean wealth change from individual gambles and selecting them on the basis of their consequences for mean final wealth.

This alternative explanation is not entirely different than the one proposed by prospect theory: Tversky and Kahneman (1992) themselves explain the nonlinear transformations of the objective payoffs and probabilities posited by prospect theory as consequences of "diminishing sensitivity" to changes in the data of the kind that are well-documented in psychophysical studies of many different sensory domains. Our goal in this paper is to further

[1]This is shown with particular clarity by experiments of Kahneman and Tversky (1979) in which the modal subject's preference between two particular probability distributions over final wealth can be reversed changing the framing of the choice problem, from a choice between a random gain and a certain gain to choice between a random loss and a certain loss. Kahneman (2002) calls the assumption that utility should be a function of wealth rather than of *changes* in wealth "Bernoulli's error."

[2]See, for example, Barretto-García *et al.* (2023), Enke and Graeber (2023), Enke and Shubatt (2023), Frydman and Jin (2022, 2023), Khaw *et al.* (2021), Netzer *et al.* (2022), Oprea (2023), Steiner and Stewart (2016), Vieider (forthcoming) and Woodford (2012). This interpretation of measured risk attitudes is consistent with an emerging literature in which behavioral anomalies that have often been treated as reflecting non-standard preferences or sub-optimal heuristics are instead attributed to optimal adaptation of decision rules to the presence of cognitive noise. See, for example, Augenblick *et al.* (2023), Azeredo da Silveira *et al.* (2020), Bhui and Xiang (2022), Charles *et al.*, (forthcoming), Enke and Graeber (2021), Enke *et al.* (2023), Gabaix and Laibson (2022), Gershman and Bhui (2020), Natenzon (2019), Vieider (2021) and Woodford (2003, 2020).

[3]Such theories need not presume that only expected financial wealth should matter for people's welfare in the case of large gambles; the presumption is simply that there should be little change in their marginal utility of additional money income in the case of the different outcomes of a small-stakes laboratory experiment. See further discussion in Khaw *et al.* (2021).

develop this insight by presenting a quantitative model of the way in which valuations of simple gambles are less than fully sensitive to changes in the data specifying the gambles, owing to random noise in the cognitive process through which intuitive judgments about the value of gambles is made. The formal structure of the theory is directly analogous to a way in which psychologists and neuroscientists often model imprecision in judgments about sensory magnitudes, which attributes it to noise in the mapping of objective magnitudes into an internal representation on some "psychological scale," or in the retrieval of that internal representation in the process of forming a judgment.[4]

Deriving insensitivity to objective conditions from deeper foundations of this kind, rather than simply positing nonlinear transducers as Tversky and Kahneman do, has several advantages. First, it allows unified modeling of the observed distortions in average valuations and the random trial-to-trial variations that must in any event be allowed for in order to fit experimental data.[5] Second, it allows us to base our quantitative specification of the imprecision in the internal representation of lotteries on empirically-based models of imprecise cognition in other domains (in particular, studies of imprecision in judgments about the size of numbers, ratios and proportions), thus unifying the study of choice under risk with other areas of cognitive psychology. Third, by explaining systematic biases in valuation as optimal responses to the presence of a particular type of cognitive noise (attributable to resource limitations of the brain), it provides an explanation for the persistence of such biases, to the extent that it is somewhat predictable that one should observe them under certain circumstances. And finally, it can provide an explanation for the failure of prospect-theoretic transducers to remain stable across conditions.[6]

Suppose that in a particular decision problem, the probability distributions of prospective wealth changes (positive or negative) that can result from the various alternatives on offer are conveyed by a vector of data $\boldsymbol{x}$. Noisy-coding models propose that the DM's decision is based not on the exact vector $\boldsymbol{x}$, but on a noisy mental representation $\boldsymbol{r}$, a draw from a conditional probability distribution $p(\boldsymbol{r}\,|\boldsymbol{x})$. An optimal decision rule is then one that chooses the option that maximizes the conditional expectation $\mathrm{E}[\Delta W\,|\boldsymbol{r}]$, where ΔW is the change in the DM's wealth as a result of the choice.[7] Such a theory can imply systematic departures from risk-neutral choice, insofar as (letting ΔW_a, ΔW_b be the random wealth changes implied by two options a and b) it is possible to have

$$\mathrm{E}[\Delta W_a\,|\boldsymbol{r}] \;>\; \mathrm{E}[\Delta W_b\,|\boldsymbol{r}]$$

more than half the time, so that a is chosen more often than not, even in a case where

$$\mathrm{E}[\Delta W_a\,|\boldsymbol{x}] \;<\; \mathrm{E}[\Delta W_b\,|\boldsymbol{x}]$$

so that choosing b would actually maximize expected wealth in this case. In addition, the theory provides a simple explanation for the "reflection effect" of Kahneman and Tversky

[4]See, for example, Green and Swets (1966), Thurstone (1959), and other references cited in Woodford (2020).

[5]We discuss quantitative specifications of prospect theory for empirical purposes in section 4.1.

[6]For further discussion of this point, see section 5.

[7]Calculation of this conditional expectation involves a prior distribution over possible true data $\boldsymbol{x}$, as well as the encoding probabilities $p(\boldsymbol{r}\,|\boldsymbol{x})$; hence optimality must be relative to a particular prior over possible decision problems that the DM may face. See the model presented in section 2, and the discussion of how the prior is specified in our analysis of our experimental data in section 2.6.

(1979), according to which the sign of the modal risk attitude flips if one changes gains into losses in a decision problem, while preserving the absolute values of all of the payoffs and their probabilities. If we assume a model of noisy coding under which monetary amounts are represented in the same (noisy) way regardless of whether they are gains or losses, then changing the signs of all of the payoffs, while leaving the data $\boldsymbol{x}$ otherwise the same, the probability distribution of values for $\mathrm{E}[\Delta W_i\,|\boldsymbol{r}]$ should remain the same (in the case of either option i), except with the sign reversed. Thus if the probability of an internal representation $\boldsymbol{r}$ such that $\mathrm{E}[\Delta W_a\,|\boldsymbol{r}] > \mathrm{E}[\Delta W_b\,|\boldsymbol{r}]$ was greater than 1/2 in the case of a problem involving random gains, the probability of the reverse inequality should also be greater than 1/2 in the case of the corresponding problem involving random losses.[8]

An important question about theories of this type concerns the stage in the decision process at which cognitive noise enters. Woodford (2012) and Khaw *et al.* (2021) suppose that the information provided about each of the possible payoffs is individually encoded with noise, and (in the case that probabilities vary from trial to trial, so that these should be represented imprecisely as well) that the information about probabilities is also (separately) encoded with noise; the assessed value of $\mathrm{E}[\Delta W_i\,|\boldsymbol{r}]$ for a given option i would thus depend in general on multiple components of the noisy internal representation $\boldsymbol{r}$. They assume, however, that the imprecision in people's choices results entirely from the posited imprecision in the internal representation of the original data supplied by the experimenter; perfectly precise calculation of the expected value of taking any option on the basis of the noisy readings of the payoffs and probabilities is assumed to be possible.[9] Enke and Graeber (2023) instead suppose that the DM has access to the precise values of the original data, but that noise enters in their computation of what these numbers imply about the expected utility gain from choosing a particular option. Thus in their account of the theory, the noisy internal representation $\boldsymbol{r}$ on which the DM's decision must be based is the noisy outcome of a process of mental calculation of expected utilities (taking the original data $\boldsymbol{x}$ as inputs), rather than noisy readings of the original data themselves.

One way to capture this difference in perspectives would be to contrast a model in which the internal representation $\boldsymbol{r}$ is a collection of noisy representations of individual pieces of information about the available options (for example, the size of one of the possible payoffs from one of the options) with one in which it is a noisy representation of the expected values of the presented options. Under this second hypothesis, $\boldsymbol{r}_i$ (the part of $\boldsymbol{r}$) that conveys information about option i) would be drawn from a distribution $p(\boldsymbol{r}_i\,|EV_i)$ that depends only on the value of $EV_i \equiv \mathrm{E}[\Delta W_i\,|\boldsymbol{x}]$ for that option. This in turn would imply that the optimal estimate $\hat{v}_i \equiv \mathrm{E}[EV_i\,|\boldsymbol{r}]$ of a given option should be drawn from a distribution $p(\hat{v}_i\,|EV_i)$ that depends only on the option's expected value. The kind of theory proposed by Khaw *et al.* (2021) instead allows the distribution $p(\hat{v}_i\,|\boldsymbol{x}_i)$ to be different for different lottery specifications $\boldsymbol{x}_i$ with the same expected value.[10]

---

[8] See Khaw *et al.* (2021) for further discussion.

[9] Note that an assumption that the DM's decision is based on noisy readings of the original data need not be understood to mean that there is error in the DM's perception of what the experimenter has told them. See further discussion in section 5, and also in Woodford (2020) and Khaw *et al.* (2021).

[10] Even if one supposes that the noise in $\boldsymbol{r}_i$ is purely random error in a report of the value of EV_i, the distribution of values for $\hat{v}_i$ could depend on elements of $\boldsymbol{x}_i$ other than EV_i alone if, as proposed by Enke and Shubatt (2023), the amount of noise in the computation of the lottery's expected value depends on the

While it might seem straightforward to test whether the distribution of subjective valuations reflected in subjects' choice behavior depends only on the expected values of the lotteries presented to them on each trial, early empirical tests of the noisy-coding hypothesis do not make this clear. In the experiment of Khaw *et al.* (2021), the lotteries used on different trials all involve a single non-zero outcome, received with a probability p that is the same on all trials, so that it is assumed that only the magnitude of the monetary payoff (which varies from trial to trial) is encoded imprecisely. In this experiment, the variation from trial to trial in the expected value $EV = pX$ is proportional to the variation in the monetary payoff X, so that a hypothesis that $\boldsymbol{r}_{risky}$ is a noisy semantic representation of EV would have identical implications to the hypothesis of Khaw *et al.* that it is a noisy semantic representation of the value of X on that trial.[11]

The experiments reported in Enke and Graeber (2023) also all involve lotteries with a single non-zero outcome, and in their case the payoff X is the same in absolute value on all trials (though it is sometimes a gain and sometimes a loss), while the probability p of receiving this payoff varies from trial to trial. Thus application of the theory of Khaw *et al.* to their case would involve a noisy internal representation only of the probability p on each trial. But since the variation from trial to trial in the expected value is proportional to the variation in p, the predicted distribution of subjective valuations conditional on the objective lottery characteristics will be the same whether we suppose that $\boldsymbol{r}$ is a noisy semantic encoding of the value of p (as proposed in a theoretical discussion in Khaw *et al.*, 2021) or a noisy semantic encoding of the value of EV on that trial.

We accordingly seek here to examine the degree to which the valuations of lotteries by experimental subjects can be explained by a model of noisy coding, when the probability p of the non-zero payoff, the magnitude $|X|$ of that payoff, and its sign (i.e., whether the lottery involves a random gain or a random loss) are all varied from trial to trial.[12] We begin by documenting the way in which changes in each of these features affects subjects' apparent risk attitudes, as measured by the ratio of their elicited willingness to pay (WTP) for the outcome of a lottery to the lottery's expected value (EV). (There would of course be no variation at all in the ratio WTP/EV across the various lotteries in the case of perfectly risk-neutral choice.)

While the issue of subjects' valuations of simple lotteries has already been much studied, there are a number of reasons for us to revisit the issue. One is that many studies elicit only a single valuation from each subject for a given lottery; it is assumed that a given lottery should have a well-defined value under a given person's preferences, and that one need only ask once to elicit it. Our theoretical framework assumes instead that responses

complexity of the calculation, which would plausibly differ across lotteries with the same expected value. However, it is hard to say what restrictions on the distribution $p(\hat{v}_i\,|\boldsymbol{x}_i)$ this hypothesis should imply, in the absence of specific theory of how complexity depends on the specification of the lottery. Here instead we test the implications of a quite specific model of the noisy coding of information about payoffs and probabilities.

[11] The same is true of the subsequent experimental studies of Frydman and Jin (2022) and Barretto-García *et al.* (2023), who also test the implications of the noisy-coding model using an experimental designs that closely follow the one in Khaw *et al.* (2021).

[12] In this way, we separately examine the roles of each of the three factors emphasized in prominent descriptive accounts: the fourfold pattern of Tversky and Kahneman (1992) depends on the size of p and the sign of the payoff, while the fourfold pattern of Markowitz (1952) depends on the sign and magnitude of the payoff.

to a given decision problem should vary randomly from trial to trial, owing to noise in the internal representation $\boldsymbol{r}$; and we are furthermore interested in measuring this randomness, because our model of optimal adaptation to cognitive noise implies that biases in the average valuation of a given lottery depend on the degree of noise in the internal representation. To test such a theory, we need to fit the predictions of our model to data on *both* average valuations and the degree of trial-to-trial variation in these valuations. Hence (as in Khaw *et al.,* 2021) we adopt an experimental procedure from studies of the accuracy of perceptual judgments, in which a variety of items are presented to the subject for evaluation in random order, with the same item (in our case, the same lottery) appearing multiple times over the course of the experiment.[13]

We also wish to separately study the effects of variation in p and variation in $|X|$ on apparent risk attitudes, varying each independently of the other, and varying each over a non-trivial range; we also consider both lotteries involving gains and ones involving losses. While the effects of varying p over a range that spans both very low and very high values have often been studied, such studies often follow Tversky and Kahneman (1992) in directing attention primarily to the way that the size of WTP/EV depends on p — a dependence that is interpreted as evidence regarding the shape of the probability weighting function of prospect theory — with relatively little attention to the effects of stake sizes. A smaller literature (e.g., Hershey and Schoemaker, 1980; Scholten and Read, 2014) has focused instead on the effects of variation in the stake size (stressed by Markowitz, 1952), but frequently considers only lotteries in which p, the probability of a non-zero payoff, takes one of a few very small values. We instead consider a wider range of values for both p and $|X|$, and in the case of lotteries involving both gains and losses.[14]

Our experimental findings are broadly consistent with those obtained by previous authors; in particular, for the median size of payoff used in our experiment, the effects of varying both p and the sign of the payoff are consistent with the fourfold pattern of Tversky and Kahneman (1992). At the same time, we find clear stake-size effects, of the type reported by previous authors:[15] for a given value of p and sign of the payoffs, increasing stake size leads to more risk-averse (or less risk-seeking) valuations in the case of lotteries involving gains, but to more risk-seeking (or less risk-averse) valuations in the case of lotteries involving losses. And like previous authors, we find the strongest effects of this kind in the case of very low values of p. A new result of ours is that the stake-size effects are roughly *log-linear*: the logarithm of WTP/EV is a decreasing affine function of the logarithm of $|X|$, though with a slope and intercept that depend on the value of p.

We show that all of these effects on measured risk attitudes can be explained — qualitatively, and to a reasonable extent even quantitatively — by a fairly parsimonious model in which

[13]An early example of the use of this method in the study of risk preferences is the work of Mosteller and Nogee (1951).

[14]A recent study by Gonzalez and Wu (2022), about which we learned only after circulation of the first draft of this paper, also elicits lottery valuations for a wide range of values of both p and $|X|$. However, the study of Gonzalez and Wu differs from ours in only considering lotteries involving gains, and in only eliciting a single valuation of each lottery from each subject. We discuss the consistency of the data of Gonzalez and Wu with our own findings in the Appendix, section G.

[15]See in particular Hershey and Schoemaker (1980), Kachelmeier and Shehata (1992), Fehr-Duda *et al.* (2010), Scholten and Read (2014), and Bouchouicha and Vieider (2017).

subjects' valuations are assumed to be optimal, subject to the presence of unavoidable cognitive noise.[16] In our model, three different types of noise — noise in the representation of monetary payoffs, noise in the representation of probabilities, and noise in response selection — have empirically distinguishable effects, and we can estimate the magnitudes of separate noise parameters to determine the importance of each type of noise in explaining our subjects' behavior. We also consider an alternative model in which we suppose that it is the lottery's EV that is retrieved with noise (i.e., that the noise arises purely in accessing the outcome of this calculation, rather than in retrieval of the values of p and X individually), and show that this type of model fits our data much less well than a model that assumes separate sources of noise in the representations of the initial data regarding payoffs and probabilities.[17]

The paper proceeds as follows. In section 1, we present the results of our new experiment. Section 2 then presents and motivates the elements of our baseline model of optimal bidding on the basis of noisy internal representations. Section 3 derives the predictions of the theoretical model for the way in which departures from risk-neutral valuations should vary with the sign of the lottery payoffs, the probability of a non-zero payoff, and the size of the non-zero payoff. Section 4 then discusses the fit of these theoretical predictions with the data moments reported in section 1, and compares the fit of our baseline model with that of a variety of alternative specifications of the cognitive noise, as well as with stochastic versions of prospect theory. Section 5 concludes with a review of additional types of evidence that, in our view, also support the interpretation of risk attitudes as reflecting adaptation to cognition noise, and discusses broader implications of the theory.

1 The Instability of Risk Attitudes: New Experimental Evidence

Here we provide additional evidence regarding the stake-dependence of risk attitudes through a new experimental study. As in previous studies following Tversky and Kahneman (1992), we elicit certainty-equivalent values for lotteries that are described to experimental subjects, and map out out the complete fourfold pattern of risk attitudes by presenting lotteries involving both gains and losses, and both large and small values of p. Our approach differs from most previous studies, however, in several respects. First, we consider a larger number of stake sizes for each value of p, in order to more precisely map the way in which the relative risk premium (the percentage difference between WTP and EV) varies with stake size. Second, we present each decision problem to the same subject many times (though not in sequence, so that subjects are unlikely to remember their previous response to the same question), because of our interest in measuring the degree of random variation in the subject's responses from trial to trial. And third, we use a different method for eliciting

[16] In concurrent work, Vieider (forthcoming) reaches a similar conclusion about binary choices between lotteries. His results complement our findings regarding elicited certainty equivalents, and provide further reason to conclude that patterns of the kind summarized by prospect theory reflect optimal adaptation to cognitive noise.

[17] On the distinguishability of different sources of error in the case of both perceptual and cognitive judgments, corresponding to different stages of the cognitive process that produces an expressed judgment, see the review by Findling and Wyart (2021).

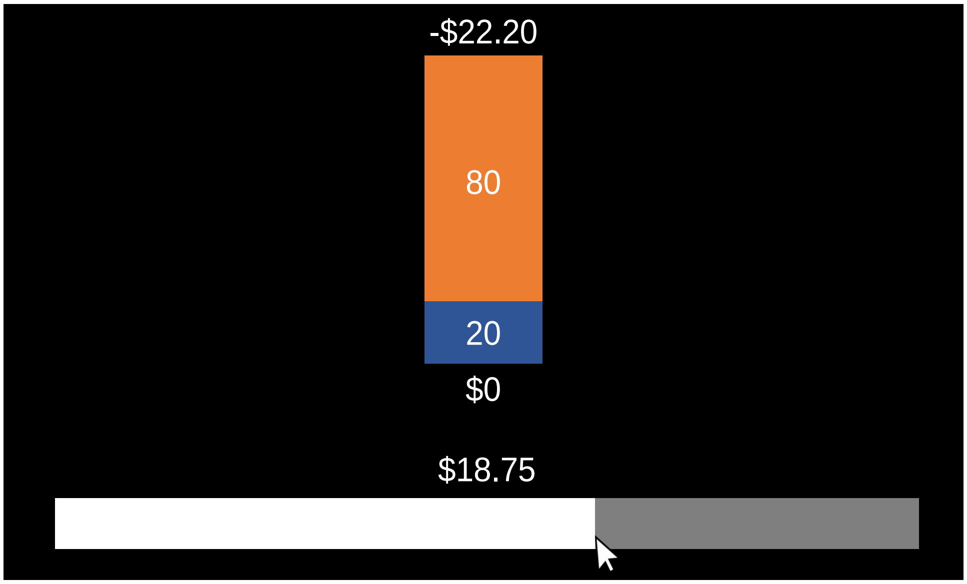

Figure 1: Example of the screen seen by an experimental subject.

subjects' valuations than in the earlier studies, also in order to make the degree of random trial-by-trial variation more visible.[18]

1.1 Experimental Design

A total of 28 subjects[19] participated in an experiment in which they were required to bid dollar amounts that they were willing to pay to obtain the outcome of a lottery which would pay an amount X with a probability p, and otherwise zero. The screen interface is shown in Figure 1. On each trial, the lottery offered is visually represented by a two-color vertical bar, the two segments of which represent the two possible outcomes. The probability of each outcome is indicated by a two-digit number inside that segment of the bar (showing the probability of that outcome in percent); the relative probabilities of the two outcomes are also indicated visually by the relative lengths of the two differently-colored segments. The monetary payoffs associated with each outcome (X and 0 respectively) are indicated by numbers at the two ends of the bar. (Note that the probabilities of *both* outcomes are displayed to the subject, with each given equal prominence, though to simplify notation we refer to the probabilities in any given decision problem by specifying only the probability of the non-zero payoff.)

A wide range of values of the probability p were used on different trials, corresponding to the different columns in Figures 2 and 3.[20] Five different values of the non-zero payoff

[18]Unlike classic early studies such as those of Hershey and Schoemaker (1980) or Tversky and Kahneman (1992), we also take care to incentivize subjects' choices, as discussed below. The importance of presenting choices involving real as opposed to merely hypothetical payoffs, especially for the measurement of stake-size effects, is illustrated by Holt and Laury (2002, 2005).

[19]These were student subjects recruited at Columbia University, following procedures approved by the Columbia Institutional Review Board under protocol IRB-AAAQ2255.

[20]The full set of 11 different probabilities were not used with any of the subjects. Instead, 13 of the

were used: \$5.55, \$7.85, \$11.10, \$15.70, and \$22.20. (These values were chosen to be roughly equal distances apart along a logarithmic scale; we did not use integer numbers of dollars, so as not to encourage subjects to treat the problem as a test of arithmetic ability.) Each of these payoffs could be either positive (a possible gain) or negative (a possible loss); thus on a given trial, X could be either \$22.20 or -\$22.20 (as in the case shown in Figure 1). Each of the possible values of p was paired with all ten of the possible values of X (both positive and negative), and the same decision problem (p, X) was presented to any given subject 8 times over the course of the experimental session, but with the problems randomly interleaved.

On each trial, after presentation of the lottery, the subject was required to indicate the amount that they were willing to pay for the outcome of the lottery, by moving a slider in a horizontal bar using the computer mouse. In the case of a lottery involving losses, the subject had to indicate the amount that they were willing to pay to *avoid* having to pay the outcome of the lottery. Thus in our discussion below, we refer to the subject's bid as WTP, their declared willingness-to-pay.[21] As shown in Figure 1, the dollar bid implied by a given slider position was shown on the screen. We used this method of elicitation of subjects' valuations, rather than the commonly used multiple-price-list procedure, because it allowed subjects to give a precise response rather than only indicating an interval. The fact that subjects' responses were not exactly the same on multiple repetitions of the same decision problem is not a disadvantage of the procedure in our case; the variability of trial-by-trial responses is actually one of the things that we wish to measure, rather than being regarded as a nuisance. Subjects' choices were incentivized by selecting one of their trials at random at the end of the experiment to be the one that mattered, and then conducting a BDM auction (Becker, DeGroot, and Marschak, 1964) in which the subject's bid on that trial was compared with a random bid chosen by the computer (independent of the subject's bid).[22]

On some trials, subjects submit a bid of zero (the leftmost position of the slider).[23] Since a subject should never be genuinely indifferent between the lottery offered and zero for sure (the lottery either clearly dominates zero, in the case of a random gain, or is clearly dominated by zero, in the case of a random loss), we interpret these responses as a subject declining to bid, rather than a considered bid that happens to be equal to zero. The trials on which the subject declines to bid are discarded in the analysis below of subjects' willingness-to-pay. (See section 2.5 for our theoretical interpretation of the zero-bid trials.)

1.2 Results

Figures 2 and 3 present statistics regarding subjects' reported willingness-to-pay (WTP) for each of 110 different lotteries: 11 different values of p (the eleven columns), and 5 different

subjects completed 400 trials each, in which five values of p were used; the other 15 subjects completed 640 trials each, in which eight different values of p were used. This allowed us to have multiple repetitions of the same problem for each of the subjects, in order to obtain a clear measure of trial-to-trial variability in the subject's response to each problem, without requiring excessively long experimental sessions. The particular values of p used with different groups of subjects are explained in the Appendix, section E.1.

[21] In the case of a lottery involving losses, we define WTP as the negative of the amount indicated by the subject's slider, so that in all cases WTP represents an elicited certainty-equivalent value of the lottery.

[22] The incentives created by this procedure are discussed further in the Appendix, section C.1.

[23] This occurs about 1.2 percent of the time overall, though more frequently when the EV of the lottery is small. See the Appendix, section E.2, for more information about these bids.

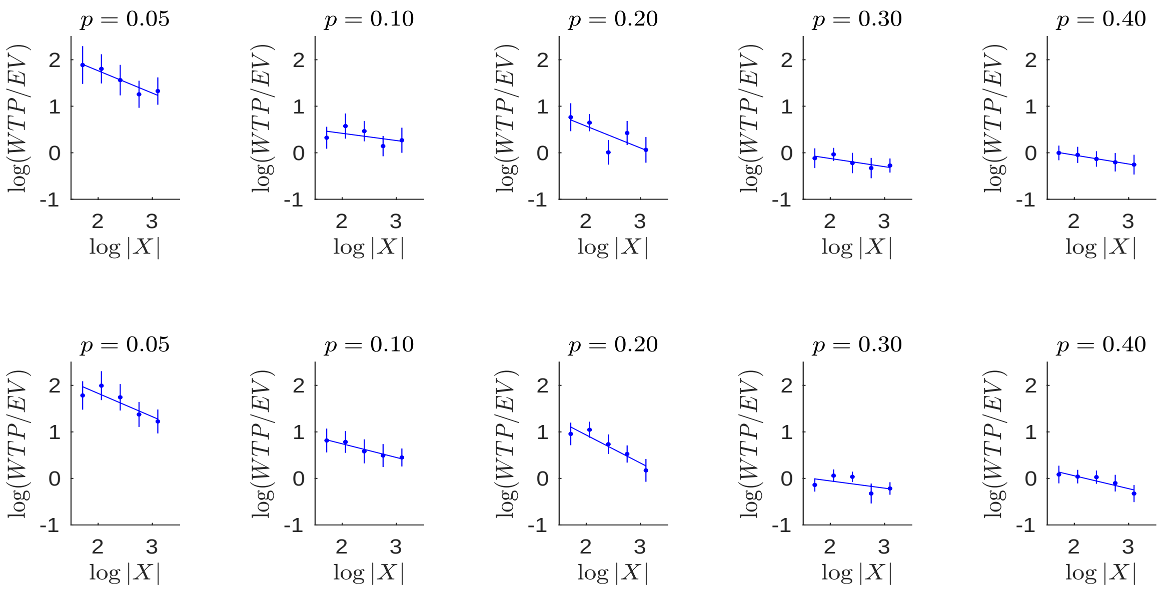

Figure 2: The distribution of values for WTP as a multiple of EV, for lotteries with different values of p (the different columns) and $|X|$ (the horizontal axis within each panel). The top panel in each column refers to lotteries involving random gains ($X > 0$), and the bottom panel to lotteries involving random losses ($X < 0$).

values of $|X|$ (the horizontal axis of each panel), in both the case of random gains (the top panel of each column) and the case of random losses (the lower panel of each column). For each lottery, subjects' bids are described in terms of the implied value of $\log(WTP/EV)$, where the expected value of the lottery is given by $EV = pX$. This can be interpreted as a measure of the relative risk premium in the case of lotteries involving random losses; the negative of this quantity measures the relative risk premium in the case of lotteries involving gains. Risk-neutral valuations (or perfectly accurate bidding, given the objective assumed in equation (2.4) below) would correspond to a value of zero on every trial, for each lottery (p, X). Thus the statistics presented in the figures measure the degree of discrepancy with respect to this benchmark, for those trials on which the subject submits a (non-zero) bid.[24]

In the case of each lottery, the dot indicates the mean value of $\log(WTP/EV)$, pooling all subjects. The vertical whiskers mark an interval $\pm s$ around the mean, where s is the standard deviation of $\log(WTP)$ for an "average" subject, computed as the mean of s.d.$[\log WTP]$ across the subjects who evaluate that lottery.[25] The horizontal line in each panel indicates the prediction of an OLS regression model (with separate coefficients for each panel). Figure

[24]Of course, the fact that subjects sometimes decline to bid is also a departure from risk-neutral optimizing behavior, but one that we do not model in this paper.

[25]We indicate the mean value of $\log(WTP/EV)$, rather than the median, because this is the quantity for which we derive a theoretical prediction below, which we wish to compare to the data moments plotted here (see Figures 5 and 6). We indicate the standard deviation for an average subject, rather than a measure of the overall variability of the pooled responses, because we wish to obtain a measure of the degree to which subjects' responses are noisy, rather than of the degree to which subjects' valuation rules may differ. The computation of the data moments for the "average subject" are discussed further in the Appendix, section D.1.

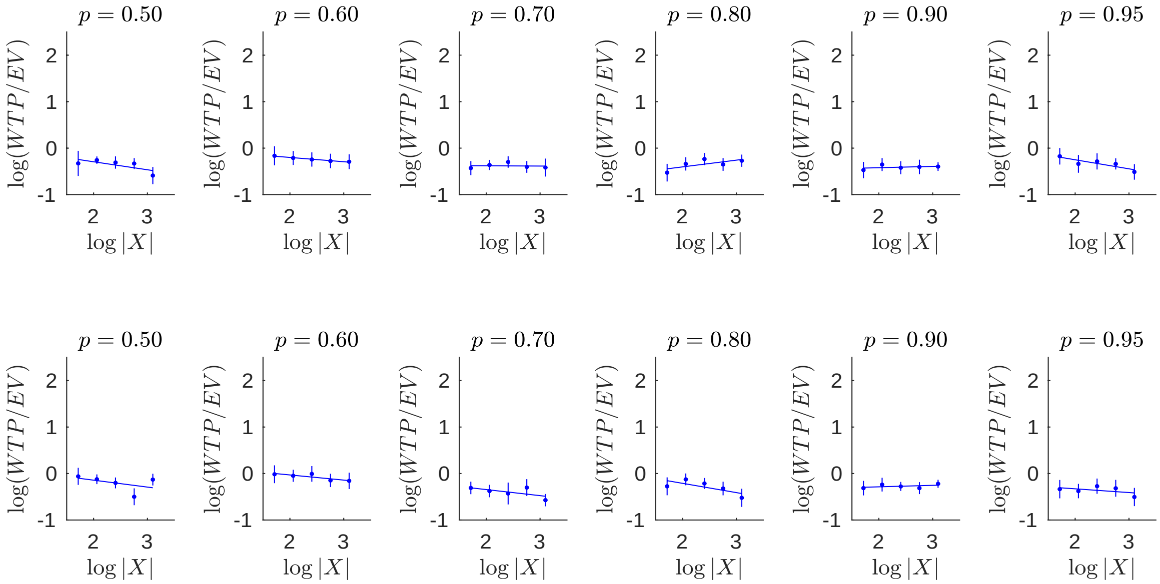

Figure 3: The same information as in Figure 2 (and using the same format), but now for probabilities $p \geq 0.50$.

2 shows the distributions of bids in the case of lotteries with relatively low values of p (between 0.05 and 0.40), while Figure 3 shows the corresponding distributions in the case of larger values of p (between 0.50 and 0.95).

Several features of our data are immediately evident from these figures. First, we see that our experiment confirms the fourfold pattern of risk attitudes documented by Tversky and Kahneman (1992): subjects' bids are for the most part risk-averse in the case of risky gains when p is 0.30 or larger ($0 < WTP < EV$), and in the case of risky losses when p is 0.10 or less ($WTP < EV < 0$), but are instead mostly risk-seeking in the case of risky gains when p is 0.10 or less ($0 < EV < WTP$), and in the case of risky losses when p is 0.30 or larger ($EV < WTP < 0$).

Yet in addition, we also see a consistent stake-size effect: in each of the 22 panels, the geometric mean value of WTP/EV becomes smaller the larger the value of $|X|$. In the transitional case (with respect to the Tversky-Kahneman pattern) where $p = 0.2$, this means that for small stake sizes we observe risk-seeking bidding in the gain domain but risk-averse bidding in the loss domain, while for larger stake sizes we instead observe risk-averse bidding in the gain domain and risk-seeking bidding in the loss domain (the alternative fourfold pattern of Scholten and Read, 2014).[26] But the sign of the stake-size effect is the same (in both the gain and the loss domains) for all of the other values of p as well, though stake-size effects are most dramatic in the case of the smallest values of p (as is consistent with the previous findings summarized in the introduction).

We also observe that the stake-size effects in each panel are approximately log-linear: the mean value of $\log(WTP/EV)$ for each lottery comes close to falling on the regression line for

[26]The same alternative fourfold pattern is observed, though in a less pronounced way, when $p = 0.4$, since in this case the mean relative risk premium changes sign for the smallest value of $|X|$. We can also observe the relative risk premium changing sign, in the direction predicted by the alternative fourfold pattern, for the cases $p = 0.1$ and $p = 0.25$ in the data of Gonzalez and Wu (2022). See Figure 9 in Appendix section G.

that panel, meaning that (fixing p and the sign of X) mean $\log(WTP/EV)$ is a decreasing linear function of $\log|X|$. Moreover, not only is the slope of this linear relationship negative (or at least non-positive), it is never more negative than -1, so that increasing the stake size (for given p) increases the mean $\log|WTP|$, as one might expect.

Finally, we observe not only that subjects do not bid in accordance with risk-neutral valuations on average; their bids for the same lottery vary from trial to trial. This within-subject variability of responses is non-trivial in the case of all of the lotteries (at least for the average subject), but it is especially notable when the probability p of the non-zero payoff is small. This is worth noting, because stake-size effects are also largest when p is small; and under the theory that we propose, it is not an accident that these two phenomena are most visible in the same cases.[27]

1.3 Log-Linear Stake-Size Effects

Visual inspection of Figures 2 and 3 suggests a downward-sloping log-linear relationship between WTP/EV and the size of the monetary payoff X in each of the panels, and moreover that this relationship is essentially the same regardless of the sign of X. Here we present statistical evidence that this is indeed an accurate characterization of our average subject's responses. We distinguish between a series of progressively more restrictive statistical models of our subjects' behavior. In the most general (purely atheoretical) characterization of the data, we suppose that for each lottery (p, X) there is a distribution of values for the willingness-to-pay of the form

$$\log \frac{WTP}{EV} \;\sim\; N(m(p,X),\, v(p,X)). \tag{1.1}$$

In what we call our "unrestricted model," there are thus two parameters, $m(p,X)$ and $v(p,X)$, to be estimated for each lottery, with no restrictions linking the parameters for any given lottery to those for any other lotteries. Our "symmetric model" instead imposes the restrictions $m(p,X) = m(p,-X)$ and $v(p,X) = v(p,-X)$, so that the distribution of values for WTP/EV depends only on p and $|X|$: it is the same for random losses as for random gains.

Alternatively, we can restrict the general model by assuming that for any p and any sign of X, $m(p,X)$ be an affine function of $\log|X|$. Our "general affine model" allows the slope and intercept for each value of p differ depending whether gains or losses are involved; this is the characterization of the data assumed in fitting the regression lines shown in each of the panels of Figures 2 and 3. Our "symmetric affine model" imposes all of the restrictions of both the symmetric model and the general affine model, so that

$$m(p,X) \;=\; \alpha_p \,+\, \beta_p \log|X|, \tag{1.2}$$

regardless of the sign of $|X|$, for coefficients (α_p, β_p) that depend only on the value of p. The "bounded symmetric affine model" imposes all of these restrictions, plus the further restriction that $-1 \,\le\, \beta_p \,\le\, 0$ for all p.

[27]It is perhaps also no accident that declining to bid at all is most common in the case of those lotteries where cognitive uncertainty is greatest, if (as our theory below assumes) greater within-subject trial-to-trial variation in bids is a sign of greater uncertainty about the value of those lotteries. However, we do not here model the decision to decline to bid.

Model	# params	LL	BIC	K
unrestricted model	220	-1503.0	3334.7	1
symmetric model	110	-1525.5	3297.9	9.8×10^7
symm. cubic	99	-1526.7	3347.6	0.0016
symm. quadratic	88	-1529.6	3312.3	73,000
general affine model	154	-1513.2	3331.0	6.4
symmetric affine model	77	-1531.7	3275.3	7.9×10^{12}
bounded symm. affine	76	-1531.8	3271.2	6.1×10^{13}
$\beta_p = 0$ for $p \geq 0.5$	71	-1534.0	3257.7	5.3×10^{16}
$\beta_p = 0$ for $p \geq 0.3$	69	-1537.0	3256.7	8.7×10^{16}
no stake effects	66	-1546.8	3264.5	1.8×10^{15}
cognitive noise model	3	-1602.5	3236.3	2.3×10^{21}

Table 1: Measures of the goodness of fit of alternative statistical models of the average subject's responses. For each model, the number of free parameters (# param), the log likelihood (LL), and the Bayes Information Criterion (BIC) are reported, as well as the Bayes factor K by which each model is preferred to the unrestricted model.

We also consider a family of models that impose even tighter restrictions on the values of the $\{\beta_p\}$. For each possible threshold p^*, we consider a model that imposes all of the restrictions of the bounded symmetric affine model, and in addition requires that $\beta_p = 0$ for all $p \geq p^*$. The most restrictive case is the "no stake effects" model that requires that $\beta_p = 0$ for all p. Consideration of more restrictive models in which β_p is required to equal zero for all large enough p allows us to obtain quantitative measures of the importance of allowing for stake effects in order to match our data.[28]

Finally, we consider models in which $m(p, X)$ is allowed to be a non-linear function of $\log |X|$, but still one with fewer free parameters than our (otherwise unrestricted) "symmetric model."[29] Specifically, we consider special cases of the symmetric model in which $m(p, X)$ is a quadratic or cubic function of $\log |X|$; these are the models called "symmetric quadratic" and "symmetric cubic" in Table 1.[30]

Table 1 reports measures of the goodness of fit of each of these models to the data on the distribution of bids of the average subject. Given that each of the models assumes a log-normal distribution of responses (1.1), the likelihood of the data under any specification of the model parameters is a function of 220 data moments: the quantities $(\hat{m}_j, \hat{v}_j)$ for each of the 110 possible lotteries (p_j, X_j). Here for each lottery j, $\hat{m}_j$ is the mean and $\hat{v}_j$ the variance

[28] We might also consider more complex patterns according to which β_p is constrained to equal zero for some but not all values of p. But Figures 2 and 3 suggest that the degree to which β_p is non-zero becomes progressively more pronounced the lower is p, and Figure 4 shows that this is true for most individual subjects as well. Our cognitive noise model, set out below, provides a reason for this to be the case.

[29] Since the BIC criterion used for model comparisons in Table 1 penalizes additional free parameters, failing to consider alternatives intermediate between linearity and a fully flexible function might allow us to accept a linear relationship as the "best fitting" one even when there is significant (but relatively simple) nonlinearity in the relationship.

[30] To reduce the size of the table, we do not also present statistics for the asymmetric variants of these models, but only the ones that assume a common relationship in both gain and loss domains.

of the sample distribution of values for $\log(WTP/EV)$. The values of these moments that we attribute to the average subject are the ones plotted in Figures 2 and 3. The likelihood also depends on N_j, the number of trials on which lottery j is evaluated. (See the Appendix, section D.1, for further details.) The parameters of each model are chosen to maximize the likelihood of these data moments, subject to the restrictions specified above.

The first column of the table reports the maximized value of the log likelihood (LL) for each model.[31] As one would expect, each successive additional restriction on the model reduces the optimized value of LL. The second column instead reports the value of the Bayes Information Criterion (BIC) for each model, defined as $BIC \equiv -2LL + \sum_k \log N_k$, where for each free parameter k of the model, N_k is the number of observations for which parameter k is relevant.[32] This is a measure of goodness of fit which (unlike LL alone) penalizes the use of additional free parameters, making it possible for a more restrictive model to be judged better (as indicated by a lower BIC). The final column provides an interpretation of the BIC differences between the different models, by reporting the implied Bayes factor K by which the model in question should be preferred to the unrestricted model (used as the baseline).[33]

While the log likelihood is lower for more restrictive versions of the model, the BIC can also be lower, if the greater parsimony of the more restrictive model outweighs the somewhat poorer fit to the individual data moments. This is what we find when we move from the unrestricted model to the bounded symmetric affine model: while LL is reduced (by 28.8 log points), the BIC nonetheless falls (by 63.5), corresponding to a Bayes factor in favor of the more parsimonious model of more than 60 trillion. This is also a lower BIC (and correspondingly a larger Bayes factor) than in the case of any of the less-restricted models, such as the general symmetric model, or the quadratic or cubic models.[34] Thus our data are more consistent with a characterization of the form assumed by the bounded symmetric affine model.

When we consider additional restrictions on the β_p coefficients, we find that the BIC can be further reduced (and the Bayes factor corresponding increased) by imposing the restriction $\beta_p = 0$ for all large enough values of p; this is illustrated in the table for the cases in which the cutoff probability is either 0.3 or 0.5, and also for the case in which we require β_p to be zero for all p. The largest Bayes factor is obtained if we set $\beta_p = 0$ for all $p \geq 0.3$. The fact that the Bayes factor is larger for this model that for the one with no stake-size effects (as indeed would also be true in the case of a higher cutoff, such as 0.5) means that we do find

[31]In all of the numerical results reported, "logarithm" means the natural logarithm. The value of LL reported here takes account not only of the likelihood of subjects' responses given the lottery (p_i, X_i) with which they are presented on each trial i, but also of the likelihood (under the estimated priors) of being presented with the sequence of lotteries $\{(p_i, X_i)\}$. The reason for including the likelihood of the lottery data under the estimated priors is to allow comparability of these LL measures with the one reported in the case of optimal bidding subject to cognitive noise (discussed further below). This definition simply adds a constant to the reported value of LL for each of the atheoretical models, so it does not affect our maximum-likelihood parameter estimates or any of the model-comparison statistics for choosing between the different possible atheoretical models.

[32]See, for example, Burnham and Anderson (2002), p. 271.

[33]The Bayes factor K in favor of model M_2 over model M_1 is given by $\log K = (1/2)[\mathrm{BIC}(M_1) - \mathrm{BIC}(M_2)]$. See Burnham and Anderson (2002), p. 303.

[34]The cubic does not even have as low a BIC as the unrestricted model: though it is more parsimonious, the reduction in the log-likelihood owing to the restrictions outweighs the reduction in the penalty for free parameters. As a result, the Bayes factor in favor of the restricted model is less than 1 in this case.

statistically significant stake-size effects, with the same sign as those reported by authors in the tradition started by Markowitz (1952) — i.e., that WTP/EV is a decreasing function of $|X|$, at least in the case of all small enough values of p.

Thus the best atheoretical characterization of our data, among those considered here, is one in which WTP/EV is a log-linear decreasing function of $|X|$, with a slope that depends on p and is most clearly negative in the case of low values of p. This relationship is essentially the same regardless of whether the lotteries involve gains or losses, and the elasticity β_p is always between 0 and -1. We show below that all of these regularities are predictions of a model of optimal bidding in the presence of cognitive noise.[35]

1.4 Heterogeneity of Subject Responses

In the previous section, we have characterized only the behavior of an average subject, by presenting for each lottery the median values of the individual subjects' mean $\log WTP$ and s.d.[$\log WTP$], and the main focus of our theoretical discussion will be provision of a quantitative model of those data moments. It is worth noting, however, that the bidding of the many of the individual subjects is at least qualitatively similar to the patterns shown in Figures 2 and 3.

We can reduce the number of statistics required to summarize the behavior of each of our subjects if, for each value of p faced by that subject, we report the coefficients (α_p, β_p) of a linear regression of the form (1.2). That is, we fit a symmetric affine model to the data for each of our 28 subjects, but allow the coefficients $\{\alpha_p, \beta_p\}$ and the residual variance v_j for each lottery to differ for each subject. The estimated regression coefficients for the different subjects are then plotted as functions of p in Figure 4. (Dashed lines connect the points corresponding to the coefficients for a given subject but for different values of p.)

While there is clearly variation in lottery valuations across subjects, we note that the general patterns of behavior identified in the data for the average subject hold also at the individual level, in most cases. In particular, we find stake-size effects ($\beta_p \neq 0$) in the case of the majority of our subjects, and in most cases we find that $-1 \leq \beta_p \leq 0$ holds (or is not clearly rejected) for all p. This is especially true in the case of the subjects who undertook 640 trials over the session; in this group β_p remains well below zero for the majority of subjects over the entire range of values for p.

We also observe a fairly consistent pattern across subjects in how both coefficients vary with p: α_p is larger (meaning a greater tendency toward risk-seeking in the gain domain and risk-aversion in the loss domain) for smaller values of p, and β_p is more negative (meaning more pronounced stake-size effects) for smaller values of p. In the next section, we discuss what our theoretical model predicts about the way in which these coefficients should vary with p.

Finally, we note a consistent pattern in the difference between the responses of subjects who evaluated different numbers of lotteries.[36] For all values of p, α_p tends to be larger, and

[35] The log likelihood and corresponding BIC statistic for that structural model are also shown in Table 1, on the bottom line. See further discussion below.

[36] As explained in the Appendix, section E.1, the two groups do not differ only in the number of questions that they were required to answer (which might have resulted in differences in the degree of fatigue or concentration). The groups also differ in the values of p used in the lotteries that they evaluated, though

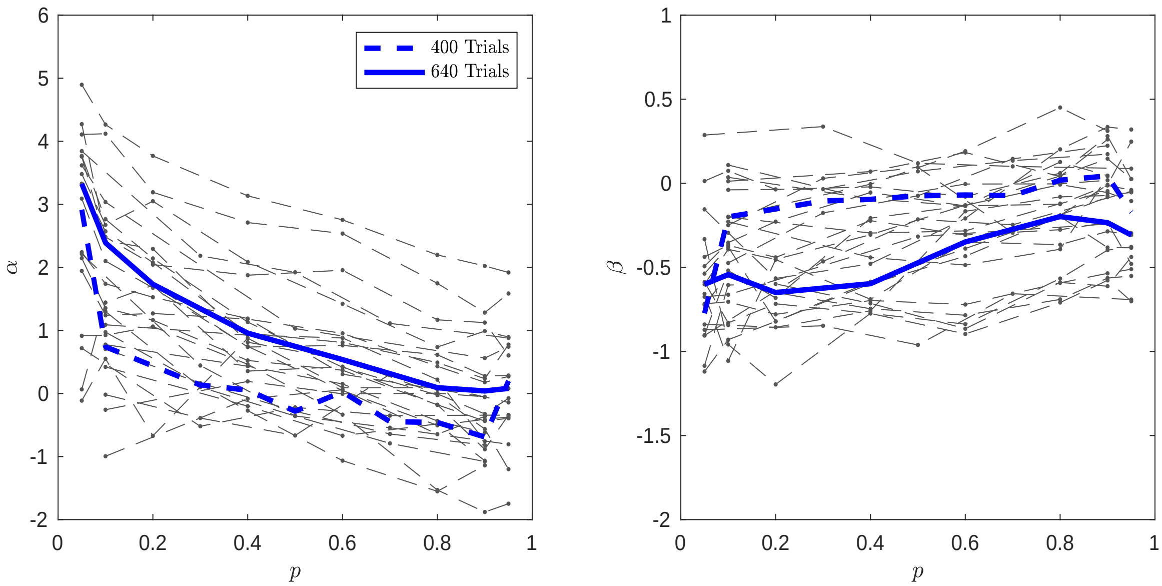

Figure 4: The coefficients $\{\alpha_p, \beta_p\}$ of the best-fitting symmetric affine model, estimated separately for each of our 28 subjects, and plotted as a function of p for each subject. The heavy curves indicate the median coefficients for each of two groups of subjects: the 13 who each completed 400 trials, and the 15 who each completed 640 trials.

β_p more negative, in the case of the subjects who undertook 640 trials relative to those who undertook only 400 trials. This suggests a possible effect of time pressure or fatigue, not simply on the variability of responses, but on a subject's average valuations. Such an effect is puzzling, if we think that subjects are reporting (though perhaps with error) valuations about which they are clear, given the specified features of the lottery; it instead has a natural explanation if we suppose that subjects' decision rules adapt in a value-maximizing way to the presence of cognitive noise, as we discuss below.

2 A Model of Endogenously Imprecise Lottery Valuation

We now show that the features of our data summarized above can be explained by a model according to which subjects' responses (on those trials in which they choose to bid) are the ones that maximize the mathematical expectation of their financial wealth, under the constraint that these responses must be based on an imprecise mental representation of the properties of the lottery that they face on a given trial, rather than upon its actual (exact) characteristics.[37] We begin by explaining our assumptions about the nature of the imprecise mental representation of the possible outcomes associated with a given lottery, and then analyze the response rule that would be optimal under the constraint that it be based on a

both groups faced both small and large values of p.

[37] The decision on some trials not to bid plainly does not maximize expected financial wealth, regardless of the imprecision of the perception of the situation on which such a decision is based. But this decision might nonetheless be an optimal adaptation if one supposes that cognitive effort can be avoided by declining to bid. See further discussion in section 2.5.

representation of this kind.

2.1 Imprecise Coding of Monetary Amounts

In our experiment, the decision problem presented on a given trial is specified by two numbers, the non-zero monetary outcome X and the probability p with which it will be received. We assume that each of these two quantities has a separate mental representation; the decision problem is mentally represented by two real numbers, r_x and r_p respectively, with r_x depending only on the value of X and r_p depending only on the value of p. We discuss first the encoding of the monetary amount, as this makes use of the same hypothesis that is explored (and tested) in our previous paper.

In Khaw *et al.* (2021), we model only the noisy coding of the monetary amount X, as the probability p is the same on all trials, and we treat the constant parameter p as understood precisely. We assume also that no mistake is made about the *sign* of X — that is, that the sign of X is encoded with perfect precision — but that the unsigned monetary amount $|X|$ is encoded probabilistically.[38] Here we again make the same assumption, and as in the previous paper, we assume that on each trial, the mental representation r_x is an independent draw from a Gaussian distribution

$$r_x \;\sim\; N(\log|X|,\, \nu_x^2(r_p)), \tag{2.1}$$

where the variance ν_x^2 may depend on r_p, the perception of how likely it is that the monetary amount will be received (and thus, how much the monetary amount matters), but is assumed to be independent of the magnitude $|X|$. In our previous paper, ν_x^2 is treated simply as a parameter (possibly differing across subjects); but it should be recalled that in our previous experiment, the probability p was the same on all trials. Since the probability varies (over a considerable range) in the current experiment, we allow for the possibility that the precision of encoding of the monetary amount may depend on it.[39]

The assumption that the mean of the distribution (2.1) grows in proportion to the logarithm of $|X|$, while the variance is independent of $|X|$, implies that the degree to which different monetary amounts can be accurately distinguished on the basis of this subjective representation satisfies "Weber's Law": the probability that a (positive) quantity X_2 would be judged larger than a quantity X_1 (also positive), on the basis of a comparison between the noisy subjective representations of the two quantities, is an increasing function of their ratio X_2/X_1, but independent of the absolute size of the two amounts.[40] There is reason to believe that the discriminability between nearby numbers decreases in approximately this way as

[38] In the experiment presented in Khaw *et al.* (2021), the monetary amounts that can be obtained are always positive. However, the paper also offers an informal discussion of how the theory can be extended to also predict choices between random and certain losses; and in that discussion it is assumed (as here) that the sign of X is encoded with perfect precision, while $|X|$ is encoded in the same way regardless of the sign of X. In the present context, this assumption is motivated by the observation in the previous section that the distribution of values for WTP/EV depends on p and $|X|$, but is (at least to a first approximation) independent of the sign of X.

[39] In our baseline model, we make a specific hypothesis about the nature of this dependence, discussed in section 2.4. We also consider an alternative model in which ν_x^2 is exogenously fixed, regardless of the probability that the lottery pays off. This variant model fits our data less well, as shown in Table 3 below.

[40] Note that (2.1) implies that the probability that $r_{x2} > r_{x1}$ is an increasing function of $\log X_2 - \log X_1$.

numbers become larger; the regularity is well-documented for numerosity perception in the case of visual or auditory stimuli (for example, judgments as to whether one field of dots contains more dots than another),[41] and there is also evidence for a similar pattern in the case of quick judgments about symbolically presented numbers, or symbolically presented numbers that must be recalled after a time delay.[42] For example, Dehaene and Marques (2002) ask subjects to recall the prices of items that they have previously been told, and find similar *percentage* errors in prices of higher- and lower-priced goods; this is consistent with a model in which what is later retrieved is a noisy semantic representation of the monetary amount that the subject had previously been told, with an error structure of the kind specified in (2.1).

2.2 Imprecise Coding of Probabilities

In our experiment, the probability p also varies from trial to trial, and must be monitored in order to decide how much to bid for a particular lottery. Hence it is natural to assume an imprecise internal representation of this information as well. We suppose that on each trial, the mental representation r_p is an independent draw from a Gaussian distribution

$$r_p \sim N\big(\log \frac{p}{1-p}, \nu_z^2\big), \tag{2.2}$$

where ν_z^2 is independent of p. The assumption that the mean of this distribution is given by the log odds of the non-zero monetary outcome means that the mean might in principle take any value on the entire real line, as in the case of our hypothesis (2.1), despite the fact that p must belong to the interval $[0, 1]$.

There are a variety of reasons for choosing the specification (2.2) for the form of the encoding noise, following a suggestion of Khaw *et al.* (2021).[43] It implies that the degree to which the relative odds of the two outcomes in two similar lotteries can be clearly distinguished depends on how different the *log odds ratio* is in the two cases; thus people should more accurately distinguish a 5 percent probability from a 10 percent probability than they distinguish a 40 percent probability from a 45 percent probability (even though there is a difference of 5 percent in each case). This assumption is consistent with the finding of Frydman and Jin (2023) that people are more accurate at judging the relative size of two small fractions (or two fractions near 1) than they are at judging the relative size of two fractions that are both close to 1/2.[44] Eckert *et al.* (2018) further find that the discriminability of two different relative frequencies ($p_1/1-p_1$ versus $p_2/1-p_2$) is increasing

This is essentially the interpretation of Weber's Law (in other sensory domains) proposed by Fechner ([1860] 1966).

[41] See Krueger (1984), and other references cited in Khaw *et al.* (2021).

[42] See Moyer and Landauer (1967), and other references discussed in Dehaene (2011).

[43] See further discussion in the Appendix, section B. In particular, we discuss there the consistency of our proposed model of imprecise representation of probabilities with evidence that Zhang and Maloney (2012) review from perceptual studies.

[44] It is also consistent with the suggestion by Tversky and Kahneman (1992), that people exhibit "diminishing marginal sensitivity" to information about probabilities as the probability moves farther from either of two "reference points," one at zero and the other at a probability of 1. Gonzalez and Wu (1999) provide further discussion and experimental evidence.

in the “ratio of ratios,” or equivalently, in the difference between the log odds in the two cases, just as our model would predict.

Our specification is also consistent with the findings of Enke and Graeber (2023), who show that subjective uncertainty about the certainty-equivalent value of lotteries like the ones in our experiment varies as an inverse-U-shaped function of the value of p (that is, higher for intermediate values of p than for either very small or very large values). If we interpret the subjective uncertainty about lottery values in their experiment as a consequence of uncertainty about the value of p implied by a given noisy internal representation r_p,[45] then this result suggests that the way in which the conditional distribution of r_p varies with p makes nearby values of p more difficult to distinguish in the case of intermediate values of p. This is in fact implied by (2.2), given the nature of the log odds transformation.[46]

2.3 Imprecise Response Selection

Our baseline model allows for a further type of cognitive noise: in addition to assuming that the DM's response must be based on noisy internal representations (r_p, r_x) rather than on the precise quantities p and X, we assume that, rather than the DM being able to choose a bid C that is a perfectly precise function of those internal representations, the response also involves inevitable imprecision. Specifically, we assume that on any given trial, the DM's response C is a monetary amount with the same sign as X (we assume no imprecision in either the DM's recognition of the sign of X, or in their recognition of the sign of an appropriate response), but a magnitude that is an independent draw from a log-normal distribution,

$$\log |C| \sim N(f(\mathbf{r}), \nu_c^2), \tag{2.3}$$

where the mean can depend on the DM's internal representation of the decision situation $\mathbf{r} \equiv (r_p, r_x, \mathrm{sign}(X))$, and the parameter ν_c measures the degree of unavoidable imprecision in the DM's response.

The randomness in (2.3) can be interpreted as resulting from random error in assessing the degree to which a contemplated symbolic response C accurately matches the DM's subjective sense of the value that should be assigned to a particular lottery. The quantity $f(\mathbf{r})$ might be taken to represent this subjective sense, which we assume to be optimally calibrated, but not to have a direct symbolic expression; the additional random error arises when the DM must decide which number symbol corresponds to this degree of value.[47] Because the

[45]See the explanation in the Appendix, section B.2, of how our model can be used to explain the results of Enke and Graeber. At least through the lens of the model of their subjects' behavior proposed in the Appendix, uncertainty about the certainty equivalent should be purely a reflection of the posterior uncertainty about p conditional on r_p.

[46]For example, if we use the Fisher information as a local measure of the discriminability of nearby probabilities on the basis of noisy internal evidence of this kind, the specification (2.2) implies a Fisher information $I \sim [p(1-p)]^{-2}$. The reciprocal of this (a local measure of uncertainty rather than of precision) is then proportional to $[p(1-p)]^2$, an inverse-U-shaped function of p, symmetric around a maximum at $p = 0.5$

[47]Similarly, in Khaw *et al.* (2021), when a DM chooses whether she would prefer a certain amount C to a lottery (p, X), there is assumed to be random error in the internal representation of the quantity C, as well as in the internal representation of what the lottery offers, and both types of randomness contribute to the stochasticity of observed choices.

additional noise relates to assessing the value of a monetary amount C proposed by the DM, it is assumed to be independent of the representation of the lottery to which the amount C is to be compared. The assumption that this noise results in a log-normal distribution for the monetary bid C is motivated in the same way as our specification (2.1) for the internal representation of the monetary amount offered by a given lottery. The distribution (2.3) implies that the probability distributions of different possible subjective valuations f that can be equated with each of two different monetary amounts C_1 and C_2 overlap to an extent that depends on the difference between $\log C_2$ and $\log C_1$, and hence on the ratio C_2/C_1, but not on the absolute magnitude of either monetary amount. This means that once again, we assume a "Weber's Law" relation for the degree of discriminability of different monetary amounts when the DM's intuitive sense of their magnitudes is consulted.

Given the unavoidable randomness specified in (2.3), we assume that the bidding intention $f(\mathbf{r}$ is optimal conditional on the internal representation $\mathbf{r}$ of the lottery currently under consideration. This means that f is chosen so as to minimize the expected loss[48]

$$\mathrm{E}[(C - pX)^2\,|\mathbf{r}], \tag{2.4}$$

when the joint distributions of $p, X, \mathbf{r}$, and C are determined by prior distributions for p and X (discussed below), and the conditional distributions specified in (2.1), (2.2), and (2.3). Note that optimality requires that the intention f (but not the DM's actual bid C) should be a deterministic function of $\mathbf{r}$.

In the special case of zero response noise ($\nu_c = 0$), the optimal bidding rule is simply

$$C \;=\; \mathrm{E}[pX\,|\mathbf{r}], \tag{2.5}$$

as assumed in our introductory discussion. The DM's response would be the mean of their posterior distribution over possible values of EV, conditional on the internal representation $\mathbf{r}$, as often assumed in Bayesian "ideal observer" models of perceptual estimates (e.g., Petzschner *et al.,* 2015; Wei and Stocker, 2015, 2017). More generally, however, this will not be true. Note that we do not, as in some models of response noise, assume that the DM chooses an intended response $f(\mathbf{r})$ that would be optimal in the absence of such noise, though the actual response differs from the intended one owing to a noise term. We instead assume that the function $f(\mathbf{r})$ is optimized for the particular degree of cognitive noise to which the DM is subject — taking into account both the encoding noise in the internal representations *and* the fact that the DM's bid will involve response noise (if $\nu_c > 0$).[49]

2.4 Endogenous Precision

In (2.1), we allow the precision ν_x^{-2} of the internal representation of the monetary amount $|X|$ on a given trial to depend on r_p, the internal representation of the probability of occurrence

[48]As we discuss in the Appendix, section C.1, the incentives provided by the BDM auction at the end of our experiment imply that the increase in the DM's expected financial wealth from bidding on a lottery is (approximately) a constant minus a positive multiple of the expression (2.4). Thus an assumption that the bidding rule minimizes (2.4) is equivalent to assuming that it maximizes the DM's expected financial wealth.

[49]If response error were assumed to be additive in (2.3), a "certainty equivalence" result would obtain: (2.5) would still be the optimal value for the intended bid, though the actual bid would equal this plus a mean-zero noise term. But because we have (more accurately, in our view) specified a multiplicative response error in (2.3), the optimal solution is more complex, as we discuss below.

of that nonzero outcome. The idea is that when the nonzero outcome is regarded as less likely to occur (on the basis of what can be inferred about this likelihood from the internal representation r_p), there should be less reason to exert mental resources in representing the nonzero outcome very precisely. The idea that encoding and/or retrieval of recently observed information can be variable in precision in this way is illustrated by a study of visual working memory by van den Berg and Ma (2018). These authors show that the accuracy with which experimental subjects can answer questions about what they were shown at various locations varies depending on the ex ante probability that the subject would be asked about a particular location, and interpret their results as reflecting endogenous variation in precision so as to economize on cognitive resources.

We now specify more precisely the nature of this dependence. We assume that greater precision of the internal representation is possible at a cost; specifically, we assume a psychic cost of representation of a monetary amount that is given by

$$\kappa(\nu_x) \;=\; \tilde{A} \cdot \nu_x^{-2}, \tag{2.6}$$

where $\tilde{A} > 0$ is a parameter indexing the cost of greater precision, in the same units as the losses (2.4) are expressed.[50] The assumption of a cost that is linear in the precision follows the model of endogenous precision in visual working memory that is fit to experimental data by van den Berg and Ma (2018).[51]

Our complete hypothesis, then, is that a precision parameter $\nu_x(r_p)$ is chosen for each possible probability representation r_p, and a subjective valuation $f(\mathbf{r})$ is chosen for each complete representation $\mathbf{r}$ of the presented lottery, so as to minimize total expected losses

$$\mathrm{E}[(C - pX)^2 \,+\, \kappa(\nu_x(r_p))], \tag{2.7}$$

where C is an independent draw from the distribution (2.3), and the expectation is over the joint distribution of p, X, r_p, r_x, and C, under the specified prior distributions.[52]

The model provides a complete specification of the predicted joint distribution of these variables, as a function of four parameters $(\mu_z, \sigma_z, \mu_x, \sigma_x)$ that specify the distribution of possible lotteries, and three additional free parameters $(\tilde{A}, \nu_z, \nu_c)$ that specify the degree of imprecision in internal representations.[53] (The latter three parameters specify the degree of imprecision in the representation of the quantities $|X|, p$, and $|C|$ respectively.) Since the former set of parameters are required to fit the distribution of values of p and X used in the experiment, only the latter three parameters are "free" parameters with which to explain subjects' responses, in the sense that we have no independent information about their values apart from what we need to assume to rationalize subjects' responses.

[50] The losses measured by (2.4) can in turn be converted into an average monetary loss in the way explained in the Appendix, section C.1.

[51] The assumption of a cost of precision that is linear in precision is also often used by economic theorists on the ground of its tractability; see, e.g., Myatt and Wallace (2012). We provide a possible cognitive process interpretation of the cost function in the Appendix, section A.

[52] The problem can be separately defined for each of the possible values of $\mathrm{sign}(X)$. Under an optimal solution, as discussed further below, the functions $\nu_x(r_p)$ and $f(r_p, r_x)$ are both independent of $\mathrm{sign}(X)$; for this reason, we have suppressed $\mathrm{sign}(X)$ as an argument of the function $\nu_x(r_p)$.

[53] Eventually, we find it convenient to express the quantitative predictions of our model as functions of an alternative set of noise parameters (A, ν_z, ν_c), where A is a function of $\tilde{A}$ and other parameters, defined below.

2.5 Declining to Bid

As already noted, on a few trials subjects submit bids of \$0, which we interpret as declining to bid on that lottery. We suppose that the DM's decision actually has two stages: a first decision whether to bid at all, followed by a second decision about which (non-zero) bid to make, only in the case that the first decision was to bid. We further suppose that the decision in each stage is optimized to serve the DM's overall objective, subject to the constraint that each decision must be made on the basis of an imprecise awareness of the precise decision problem that is faced on that trial. In such a two-stage analysis, one of the benefits of deciding in the first stage not to bid will be avoidance of the cognitive costs associated with having to decide what bid to make in the second stage.[54] The cognitive costs associated with undertaking a second-stage decision should include the cost $\kappa(\nu_x)$ of encoding (or retrieving) the magnitude of the monetary payoff with a certain degree of precision, but they could include other costs as well, that have not been specified above because they do not affect our calculation of the optimal second-stage bidding rule. One piece of evidence in support of the view that a zero bid avoids cognitive costs is our observation that subjects respond more quickly on average on the zero-bid trials.[55]

In this paper, we model only the "second-stage" problem, i.e., how the subjects bid on those trials where they choose to make a non-zero bid. This is done taking as given the probability that the DM will find themselves having to choose a non-zero bid in the case of a particular lottery (p, X), as a consequence of the first-stage decision rule.[56] The prior distribution that is relevant for the "second-stage" problem modeled above (specified mathematically by (2.2) and (2.1)) is not the frequency distribution with which the experimenters present different lotteries (p, X), but rather the frequency distribution with which the different lotteries become the object of a second-stage decision. This depends both on the distribution of lotteries chosen by the experimenter and on the first-stage decision rule. However, in our quantitative evaluation of the model below, we fit the parameters of the assumed prior distribution to the *empirical* frequency with which non-zero bids are made on different lotteries (p, X), and not to the frequency distribution of lotteries chosen by the experimenters. Given this, it is not necessary for us to model the DM's first-stage decision in order to derive quantitative predictions from our model of the second-stage decision.

2.6 Priors for the Optimal Adaptation Problem

The objective (2.7) that the DM's cognitive processing is assumed to minimize depends on the prior distributions from which the parameters (p, X) specifying the decision problem are expected to be drawn. In our numerical work here, we assume that regardless of the sign of

[54] A two-stage decision of this kind is completely modeled in Khaw *et al.* (2017); in that application, subjects are modeled as first deciding on each trial whether to adjust their existing response variable or not, and then (only if the outcome of the first decision was to adjust) deciding exactly what size of adjustment to make. Both decisions are modeled as made optimally subject to an information constraint; it is optimal not to adjust on all trials, because the cognitive costs associated with the second-stage decision can be avoided by opting in the first stage not to adjust.

[55] The average response time (RT) on the 175 trials on which zero bids were submitted was 2.61 seconds, while the average RT on the other trials was 3.83 seconds (nearly 50 percent longer).

[56] See the Appendix, section E.2, for further discussion.

X, the prior distribution for possible values of $|X|$ is of the form

$$\log |X| \sim N(\mu_x, \sigma_x^2), \tag{2.8}$$

for some parameters μ_x, σ_x. Apart from being mathematically convenient and parsimoniously parameterized, a prior of this form is found to fit the behavior of most subjects fairly well in Khaw *et al.* (2021).[57]

The prior distribution for p is assumed instead to be of the form

$$\log \frac{p}{1-p} \sim \text{Uniform}\,[\mu_z - \sqrt{3}\sigma_z,\ \mu_z + \sqrt{3}\sigma_z], \tag{2.9}$$

for some parameters μ_z, σ_z, which again indicate the mean and standard deviation of the prior.[58] Also, under the prior p and $|X|$ are distributed independently of one another (as is true in our experiment); and the joint distribution of $(p, |X|)$ is the same regardless of the sign of X (as is also true in our experiment).[59]

Note that in our theoretical analysis below, this assumed to be the joint distribution from which $(p, |X|)$ are drawn *conditional on the DM having decided to bid.* Because the probability of subjects' declining to bid is not independent of the values of p and X on that trial, the prior under which (2.7) is hypothesized to be minimized is the distribution of lottery characteristics conditional on the DM bidding, rather than the distribution of lotteries presented by the experimenter. In our numerical analysis of the model's predictions, we estimate the values of the parameters $(\mu_x, \sigma_x, \mu_z, \sigma_z)$ for the set of lotteries on which non-zero bids are made, rather than using the parameters for the distribution from which the lotteries in the experiment were drawn.

3 Consequences of Optimal Adaptation to Cognitive Noise

Here we derive the predictions of the model in section 2 for the data moments displayed in Figures 2 and 3. Note that we are interested simultaneously in explaining the observed biases (systematic differences between average WTP and the actual EV of the lottery) and the variability of the valuations of a given lottery. According to our theory, these two aspects of the data should be intimately connected; in the absence of random noise (the case in which $\tilde{A} = \nu_z = \nu_c = 0$), our model predicts that we should observe $WTP = EV$ on each trial. Hence the same small set of parameters must explain both features of the data.

[57] See also the replication of that work by Barretto-García *et al.* (2023).

[58] A truncated uniform distribution better fits the set of values for the odds ratio used in our experiment than a Gaussian distribution would. Note, however, that we do not literally sample the values used from a uniform distribution; only a discrete set of values of p are used, as shown in Figures 2 and 3.

[59] We need not specify a prior probability of encountering one sign of X or the other, since this variable is assumed to be known with perfect precision, and no issue of Bayesian decoding of an imprecise representation arises.

3.1 Implications of Logarithmic Encoding of Monetary Payoffs

We begin with a set of predictions that follow from the specification (2.1) for the noisy internal representation of monetary payoffs, the specification (2.3) for the errors in response selection, and the specification (2.8) for the distribution of payoff values under the prior for which the DM's bidding rule $f(\mathbf{r})$ is optimized. These predictions are independent of what we assume about the internal representation of probabilities, the prior over probabilities, or the way in which the precision with which monetary payoffs are encoded may depend on r_p. They do, however, depend on our also assuming that the bidding rule is optimized to minimize mean squared error under the prior.

Under these assumptions, the posterior distribution for $|X|$ conditional on the internal representation $\mathbf{r}$ will be log-normal, and the joint distribution of $(\log|X|, \log|C|)$ conditional on $\mathbf{r}$ will be bivariate normal. The algebra of log-normal distributions allows us to show that the Bayesian posterior mean estimate of the magnitude $|X|$ will be of the form

$$\mathrm{E}[|X|\,|\mathbf{r}] \;=\; \exp((1-\gamma_x(r_p))\bar{\mu}_x \;+\; \gamma_x(r_p)\cdot r_x), \tag{3.1}$$

where

$$\gamma_x(r_p) \;\equiv\; \frac{\sigma_x^2}{\sigma_x^2+\nu_x^2(r_p)} \tag{3.2}$$

is a quantity satisfying $0<\gamma_x(r_p)<1$, that depends on the degree of precision with which $|X|$ is encoded in the case of that value of r_p, and

$$\bar{\mu}_x \;\equiv\; \mu_x \;+\; \frac{1}{2}\sigma_x^2$$

is the logarithm of the prior mean of $|X|$. In the case of perfectly precise encoding, $\gamma_x = 1$, and the mean estimate of $|X|$ is exactly the true value of $|X|$; in the limit of extremely imprecise encoding ($\nu_x^2 \to \infty$), $\gamma_x \to 0$, and the mean estimate approaches the the prior mean $\exp(\bar{\mu}_x)$, regardless of the noisy internal representation r_x.

The optimal bidding rule can then be shown to be[60]

$$f(\mathrm{r}) \;=\; \log\mathrm{E}[p\,|r_p] \;+\; (1-\gamma_x(r_p))\bar{\mu}_x \;+\; \gamma_x(r_p)r_x \;-\; \frac{3}{2}\nu_c^2. \tag{3.3}$$

This has a fairly simple interpretation. In the absence of response noise, the optimal Bayesian decision rule would be $f=\log\mathrm{E}[pX\,|\mathbf{r}]$, and the latter quantity can be written as the sum of the logarithm of the posterior mean estimate of p (given r_p) and the logarithm of the posterior mean estimate of $|X|$, given by (3.1). In the case of response noise, the median bid is shaded downward (in absolute size) by a constant percentage that depends on the value of ν_c^2, to take account of the multiplicative error in bidding.

This rule, together with (C.4), and the encoding rules that specify the distribution of $\mathbf{r}$ for a given lottery, can then be used to predict the distribution of values for the ratio WTP/EV for each lottery. Note in particular that regardless of what we assume about the internal representation of the probability p, and about the way in which ν_x^2 depends on r_p, the model implies that

$$\mathrm{E}[\log(WTP/EV)\,|p,X] \;=\; \alpha_p \;+\; \beta_p\log|X|, \tag{3.4}$$

[60] See the Appendix, section C.2, for details of the calculation.

for some coefficients α_p, β_p that can depend on p. These coefficients should be the same regardless of the sign of $|X|$, so that the plots in the upper and lower rows of Figures 2 and 3 should look the same, as to a large extent they do.[61]

The model also implies that the mean value of $\log(WTP/EV)$ should be an affine function of $\log|X|$, with a negative slope, satisfying the bounds $-1 < \beta_p < 0$. Specifically, the predicted slope is given by

$$\beta_p \;=\; -(1-\gamma_p), \tag{3.5}$$

where γ_p is the mean value of $\gamma_x(r_p)$, averaging over the distribution of internal representations r_p associated with a particular true probability p.) This negative (but boundedly negative) slope is also what we observe in Figures 2 and 3, for all values of p.

Finally, the model implies that the log-linear relationship (3.4) should hold no matter how large the variations in $\log|X|$ may be. In our experiment, $|X|$ varies only by a factor of 4 between the smallest and largest values used in the experiment; as a result, the sign of the mean relative risk premium is independent of $|X|$, in each of the panels of Figures 2 and 3. However, our theoretical model implies that if a wider range of values of $|X|$ were used, the sign of the relative risk premium should be different for very small $|X|$ and very large $|X|$. This should be true in principle for all values of p, but it should be particularly easy to observe the sign change in the case of small p (since these are the cases in which β_p is most negative, for reasons discussed below). Thus our model also predicts that in the case of lotteries in which the probability of a non-zero outcome is small, if X is varied over a wide enough range, one should observe a positive relative risk premium (risk-averse valuations) in the case of a large enough potential gain, or a small enough potential loss, but should observe a negative relative risk premium (risk-seeking valuations) in the case of a small enough gain or a large enough loss. Thus our model predicts the alternative fourfold pattern documented by Hershey and Schoemaker (1980) and Scholten and Read (2014).

3.2 The Optimal Precision of Magnitude Encoding

We have derived above the optimal log-normal distribution of bids C in the case of internal representations (r_p, r_x), in the case of any given assumption about the precision of encoding of information about both p and $|X|$, including an arbitrary assumption about how ν_x^2 may depend on r_p. We now consider how an efficient coding system, subject to a linear cost of precision of the kind proposed above, would actually require the precision of magnitude encoding to vary with r_p. This allows us to determine how the coefficients (α_p, β_p) in (3.4) should depend on p.

Under any assumption about the function $\nu_x^2(r_p)$, we can compute the Bayesian posterior over possible decision problems (p, X) conditional on a given internal representation (r_p, r_x). Given this together with the distribution of bids implied by (3.3), we obtain a joint distribution for (p, X, C) conditional on the internal representation, and hence a conditional distribution for the value of the quadratic loss $(C - pX)^2$. This allows us to compute the conditional expectation (2.4).

[61] It is not only the coefficients α_p and β_p that should be the same; the model implies that the entire distribution of WTP/EV should be the same function of p and $|X|$, regardless of the sign of X.

Integrating this over possible realizations of r_x (for a given value of r_p), we obtain an expression of the form[62]

$$\mathrm{E}[(C - pX)^2 \,|r_p] \;=\; Z(r_p) \;-\; \Gamma\varphi(r_p) \cdot \exp(\gamma_x(r_p)\sigma_x^2), \tag{3.6}$$

where $\Gamma > 0$ is a constant; the functions $Z(r_p), \varphi(r_p)$ are each positive-valued, and defined independently of the choice of $\nu_x^2(r_p)$; and the function $\gamma_x(r_p)$ depends on $\nu_x^2(r_p)$ in the way indicated in (3.2). Thus equation (3.6) makes explicit the way in which the expected loss conditional on a given value of r_p depends on the choice of $\nu_x^2(r_p)$. We see that the expected loss is a decreasing function of $\gamma_x(r_p)$, and hence an increasing function of the choice of $\nu_x^2(r_p)$. If there were no cost of precision, it would be optimal to choose $\nu_x^2(r_p)$ as small as possible, for each value of r_p.

Taking into account the cost of precision (2.6), we instead want to choose $\nu_x^2(r_p)$ to minimize the total loss

$$\mathrm{E}[(C - pX)^2 \,|r_p] \;+\; \kappa(\nu_x(r_p)) \tag{3.7}$$

associated with the internal representation r_p. (The objective (2.7) stated above is just the expectation of this over all possible values of r_p.) Since $\gamma_x(r_p)$ is a monotonic function of $\nu_x^2(r_p)$, we can alternatively write the objective (3.7) as a function of $\gamma_x(r_p)$; let this function be denoted $F(\gamma_x(r_p); r_p)$. We can then express our problem as the choice of $\gamma_x(r_p)$ to minimize $F(\gamma_x(r_p); r_p)$.

We show in the Appendix, section C.3, that under the assumption that $\sigma_x^2 \leq 2$,[63] the solution to this optimization problem can be simply characterized. If

$$\varphi(r_p) \;\leq\; A \;\equiv\; \frac{\tilde{A}}{\sigma_x^4}\exp(\nu_c^2),$$

then the solution is $\gamma_x(r_p) = 0$, meaning zero-precision representation of the payoff magnitudes. (In this case the optimal decision rule is based on the prior distribution from which $|X|$ is expected to be drawn, but no information about the value of $|X|$ on an individual trial.) If instead $\varphi(r_p) \;>\; A$, the optimal γ_x is given by the unique solution to the first-order condition

$$\frac{A}{(1-\gamma_x)^2} \;=\; \varphi(r_p)\exp(\gamma_x\sigma_x^2). \tag{3.8}$$

Equation (3.8) has a unique solution $0 < \gamma_x(r_p) < 1$ for any r_p such that $\varphi(r_p) \;>\; A$; and this solution depends only on the value of $\varphi(r_p)$. We further show that $\gamma_x(r_p)$ is an increasing function of $\varphi(r_p)$, so that the implied value of $\nu_x^2(r_p)$ is a monotonically decreasing function of $\varphi(r_p)$, with $\nu_x^2(r_p) \to 0$ as $\varphi(r_p)$ is made unboundedly large, and $\nu_x^2(r_p) \to \infty$ as $\varphi(r_p) \to A$ from above.

These results make use of a specific assumption (2.6) about the cost of precision in magnitude encoding, but are independent of any special assumption about the way in which information about relative probabilities is encoded. Let us further suppose that the prior over

[62] See the Appendix, section C.3, for details of the derivation.

[63] This is the case of interest in our application. In our experiment, the variance of $\log |X|$ is approximately 0.26; thus a prior roughly consistent with the actual distribution of magnitudes used in the experiment would have to have a value of σ_x^2 much less than 2.

relative probabilities and the conditional distributions $r_p|p$ satisfy the following conditions: (i) the median of the distribution $r_p|p$ is an increasing function of p; and (ii) the posterior mean $\mathrm{E}[p\,|r_p]$ is an increasing function of r_p. Then since $\varphi(r_p) \equiv \mathrm{E}[p\,|r_p]^2$, the median value of $\varphi(r_p)$ will be an increasing function of p. It then follows from our results above that the median value of $\gamma_x(r_p)$ will be a non-decreasing function of p, and strictly increasing for p in the range for which the median value of r_p satisfies $\mathrm{E}[p\,|r_p] > \sqrt{A}$.

This in turn means that the median value of $\nu_x^2(r_p)$ will be a decreasing function of p, for all p large enough for the median optimal $\nu_x^2(r_p)$ to remain finite.[64] Thus the model predicts that the precision of encoding of the monetary payoff magnitude should be less, on average, the smaller the probability p that the lottery's non-zero payoff would be received. Essentially, the increasing cost of greater precision implies that it is not worthwhile to encode (or retrieve) the value of $|X|$ with the same degree of precision when the probability of that outcome being the relevant one is smaller.

This dependence of the precision of magnitude encoding on the value of p has implications for the predicted degree of trial-to-trial variability in subjects' bids for different values of p; but it also has implications for the degree of bias in their mean or median valuations of a given lottery. It follows from (3.5) that if γ_x is lower on average for smaller values of p, then β_p should be more negative the smaller is p. Hence stake-size effects should be strongest in the case of the smallest values of p, as found in our experiment and the other studies cited in the introduction.

The model also makes quantitative predictions about the way in which the intercepts of the regression lines shown in the various panels of Figures 2 and 3 should vary with p. If we measure the intercept by the predicted height of the regression line at a value of $|X|$ equal to its prior mean, we obtain

$$\alpha_p \,+\, \beta_p \log \mathrm{E}[|X|] \;=\; \mathrm{E}\big[\log \mathrm{E}[p\,|r_p] - \log p\;|p\big] \;-\; \frac{3}{2}\nu_c^2. \tag{3.9}$$

In general, this will vary with p, though the way in which the intercept depends on p depends only on the joint distribution of (p, r_p) — thus on the prior over p and the conditional distributions $r_p|p$ — and not on any aspects of the way in which $|X|$ is encoded. In the absence of any noise in the encoding of p (though an arbitrary degree of imprecision in the internal representation of $|X|$), (3.9) implies that the intercept will be a constant, the same for all p.[65] When p is instead encoded with noise, the posterior mean estimate $\mathrm{E}[p\,|r_p]$ will be subject to "regression bias," as a result of which the posterior mean estimate will mostly be larger than the true p when p is low, and smaller than the true p when p is high.[66] It then follows that when $|X| = \mathrm{E}[|X|]$, the sign of the intercept (3.9) should vary with p in the way required for the "fourfold pattern" of risk attitudes of Tversky and Kahneman (1992).

Our model therefore explains the existence of Tversky and Kahneman's fourfold pattern, if we vary p and the sign of X while maintaining a value of $|X|$ equal to the prior mean. At the same time, our model also predicts the existence of stake-size effects ($\beta_p < 0$). This means that for any value of p and either sign of X, varying $|X|$ over a sufficiently large range

[64] In the estimated numerical model discussed below, this is true for all of the values $p \geq 0.05$ used in our experiment.

[65] This constant would furthermore be zero in the absence of response noise.

[66] See the Appendix, section B.2, for further discussion of these predicted biases in probability estimation.

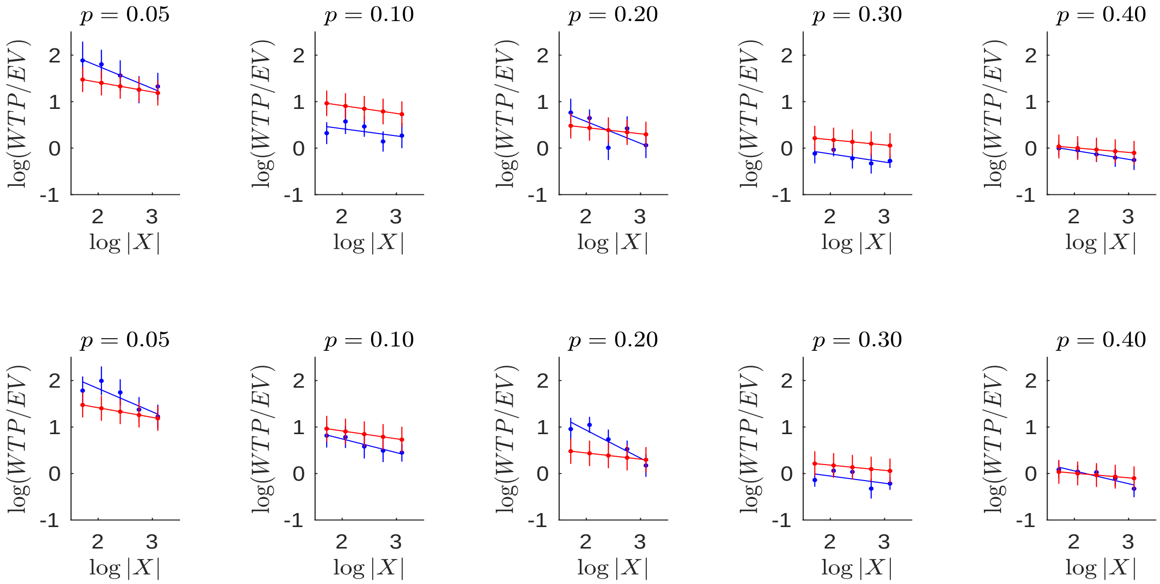

Figure 5: The same data as in Figure 2, but now compared with the predictions of the optimal bidding model with maximum-likelihood parameter estimates. (Blue: data for the average subject. Red: theoretical predictions.)

should allow one to flip the sign of the DM's relative risk premium, in a way consistent with the alternative fourfold pattern of Scholten and Read (2014). (This should be most easily visible when p is small.) Thus our model is (at least qualitiatively) consistent with both of the patterns documented in the previous literature.

4 Assessing the Quantitative Fit of the Cognitive Noise Model

We have already discussed above the statistical evidence in favor of some of the key predictions of the cognitive noise model: the prediction that the mean value of $\log(WTP/EV)$ should be an affine function of $\log|X|$ for any value of p (and the same function for lotteries involving either gains and losses), with a slope between 0 and -1. We show in Table 1 that an otherwise unrestricted "bounded symmetric affine model" provides a superior characterization of the bids of our average subject, if we assess the fit of alternative models on the basis of a BIC statistic that penalizes additional free parameters. We now consider the conformity of our average-subject data with the more detailed quantitative predictions of the model.

We test the complete set of predictions of the model set out above by finding the values of the three free parameters A, ν_z, and ν_c that maximize the likelihood of the data moments. As in our atheoretical modeling of the data in section 1, we express the likelihood of the experimental data as a function of the mean $m(p_j, X_j)$ and variance $v(p_j, X_j)$ of the distribution of bids for each lottery j specified by characteristics (p_j, X_j), and the number of trials N_j on which that lottery is evaluated. This amounts to approximating the predicted distribution of bids for any lottery, as a function of the model parameters, by a log-normal

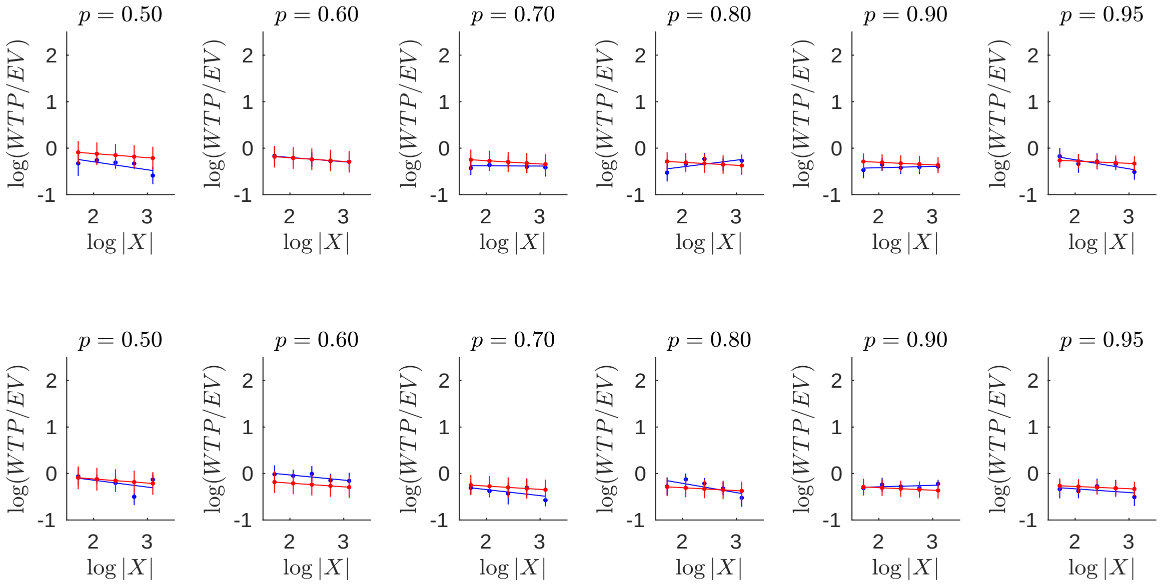

Figure 6: Continuation of Figure 5 for probabilities $p \geq 0.50$.

distribution.[67]

The theoretical data moments predicted by our model depend not only on the parameters (A, ν_z, ν_c) specifying the degree of cognitive imprecision on the part of the DM,[68] but also on the parameters $(\mu_z, \sigma_z, \mu_x, \sigma_x)$ specifying the prior distribution over possible lotteries. Thus we estimate values for all seven parameters, so as to maximize a complete likelihood function of the data, taking into account both the likelihood of the lottery characteristics presented on the different trials (under a given parameterization of the prior) and the likelihood of the subjects' bids on those trials (given our model of noisy internal representations and optimal bidding).

Figures 5 and 6 (presented using the same format as in Figures 2 and 3) show to what extent the predicted moments match the "average subject" moments when the parameters are chosen to maximize the (approximate) likelihood function.[69] The fit is not as good as that of the best-fitting affine model, shown in Figures 2 and 3; the maximized log-likelihood is a good deal lower, as shown on the bottom line of Table 1. However, the optimizing model has many fewer free parameters than the atheoretical affine model, and the BIC associated with the optimizing model is much lower than that of the affine model, as is also shown on the bottom line of Table 1. In fact, the BIC of the optimizing model is well below that of the best-fitting of the atheoretical models discussed above, namely the restricted version of

[67]We assumed such a log-normal distribution (1.1) in the case of our atheoretical data characterizations. This must be regarded as only an approximation in the case of our model of optimal bidding on the basis of noisy internal representations. For while the optimal bidding model implies a log-normal distribution of bids corresponding to each possible internal representation $\mathbf{r}$, there is a probability distribution over representations $\mathbf{r}$ for any lottery j, so that the overall distribution of bids will in general not be exactly log-normal. Our log-normal approximation is discussed further in the Appendix, section D.1.

[68]Note that the composite parameter A, rather than the quantity $\tilde{A}$ appearing in (2.6), is the measure of the cost of precision in the encoding of numerical magnitudes that can be inferred from our behavioral data.

[69]The maximum-likelihood parameter estimates for the cognitive noise parameters are shown on the bottom line of the upper part of Table 4.

Stochastic Prospect Theory			
value function	prob. weighting	LL	BIC
power law	linear	-1878.3	3775.4
power law	TK92	-1653.9	3332.9
power law	Prelec	-1627.1	3285.4
logarithmic	TK92	-1654.3	3333.6
logarithmic	Prelec	-1626.1	3283.5
Cognitive Noise Model			
baseline model		-1602.5	3236.3

Table 2: Model comparison statistics for the fit of several stochastic versions of prospect theory to the distributions of bids of our average subject. The bottom line reproduces the corresponding statistics for our baseline model, for comparison.

the bounded symmetric affine model (with $\beta_p = 0$ for all $p \geq 0.3$). The Bayes factor for the optimizing model is correspondingly larger (indeed, larger by a factor greater than 10^{21}).

4.1 Comparison with the Fit of Prospect Theory

As a benchmark for judging the degree of fit of our cognitive noise model, it is useful to compare the fit to our data of another kind of parametric model (albeit without a foundation in optimization), namely prospect theory (PT). As is well known, PT provides an explanation for the fourfold pattern of risk attitudes documented by Kahneman and Tversky, and it can be specified so as to allow for stake-size effects as well. Like our baseline model, some quantitative versions of PT involve as few as three free parameters: one to specify the degree of nonlinearity of the "value function" applied to gains or losses, one to specify the degree of nonlinearity of the "weighting function" that modifies the probabilities of the different outcomes, and one to specify the degree of random error in subjects' individual responses (Stott, 2006).

Table 2 reports the log likelihood of the average-subject data, and the corresponding BIC statistic, for several possible stochastic versions of PT, using parametric specifications of the value function and weighting function that have been popular in the empirical literature.[70] In each case, we make PT stochastic (allowing us to calculate a likelihood for our experimental data) by assuming a multiplicative response error (2.3), just as in our baseline model; but now the bidding intention $f(\mathbf{r})$ is replaced by the valuation of the lottery (p, X) implied by a deterministic version of PT. In all of the versions of PT considered in Table 2, we assume (for the sake of parsimony) that the same value function and weighting function apply in both the gain and loss domains.[71]

[70]The details of each of these specifications, and the best-fitting parameter values in each case, are explained in the Appendix, section F.

[71]Tversky and Kahneman (1992) instead estimate separate parameters for lotteries involving gains or losses. We have also estimated asymmetric versions of each of the variants of PT considered in Table 2, but find that they fit less well than the corresponding symmetric theories, according to the BIC criterion (which penalizes additional free parameters). Hence we only report our estimates for the symmetric versions of the models here.

We consider two possible specifications of the value function: the "power law" specification used by Tversky and Kahneman (1992), also extensively used in the subsequent literature, and a "logarithmic" specification advocated by authors such as Bouchouicha and Vieider (2017) as a way of matching empirically observed stake-size effects.[72] We consider three possible specifications of the weighting function. Our "linear" specification (in which the weight $w(p)$ is simply equal to p, corresponds to the expected-utility model of von Neumann and Morgenstern,[73] and has no free parameters. We also consider the one-parameter family of nonlinear weighting functions proposed by Tversky and Kahneman (1992), called "TK92" in the table, and a two-parameter family of functions subsequently proposed (on axiomatic grounds) by Prelec (1998).

Comparison of the second line of Table 2 with the first shows that allowing for nonlinear probability weighting of the kind proposed by Tversky and Kahneman (1992) improves the fit of the model enough to more than offset the penalty for the additional free parameter. The predictions of the best-fitting model using the functional forms proposed by Tversky and Kahneman (1992) are illustrated in row (a) of Figure 7.[74] The model predicts no stake-size effects of the kind observed in our data, though the nonlinearity of the weighting function allows the model to capture the fact that $\log(WTP/EV)$ is on average higher in the case of lower values of p. Comparison of the third line of Table 2 with the second shows that the more complex Prelec specification of the weighting function fits even better, again even allowing for the penalty for additional free parameters. Comparison of the fourth line with the second shows that a logarithmic value function does not fit better than the "power law" form used in Tversky and Kahneman (1992) when it is combined with Tversky and Kahneman's specification of the weighting function.[75] But when combined with the Prelec weighting function, the logarithmic value function does fit better, and in fact, the version of PT that combines a logarithmic value function with Prelec's weighting function fits our data best, according to the BIC criterion. (The predictions of this version of PT are illustrated in row (b) of Figure 7.) Finally, the bottom line of the table shows that the cognitive noise model fits better than any of the versions of PT considered here. Indeed, the difference in BIC statistics between even the best-fitting version of PT and the cognitive noise model implies a Bayes factor larger than 17 billion in favor of the cognitive noise model.

We do not mean to imply that on the basis of a single (rather small) dataset we can establish that our model is more descriptively accurate than prospect theory; our point

[72] As explained in the Appendix, section F, the "power law" specification implies that the ratio WTP/EV should depend on p, but be independent of the stake size $|X|$; and indeed, Tversky and Kahneman (1992) minimize the importance of stake-size effects in their data.

[73] This is technically not an example of prospect theory, but is formally a special case of the other models considered in the table. We include the first line of the table because this case — expected utility maximization with a CRRA utility function applied to the gains from an individual gamble rather than to total wealth — corresponds to a popular model in the experimental literature on choice under risk.

[74] In Figure 7, we show only one row for each model, as we consider here only symmetric models, in which the predictions are identical for lotteries involving gains or losses.

[75] As shown in the Appendix, section F, the best-fitting version of the logarithmic value function in this case is close to being linear; imposing instead the constraint that the value function be linear would yield a lower BIC, because of the elimination of a free parameter. In the reported results on line 2 of Table 2, where a power-law value function is assumed, we obtain a corner solution (a linear value function) as the best-fitting value function; in this case, because of the corner solution there is no free parameter to penalize in the BIC statistic.

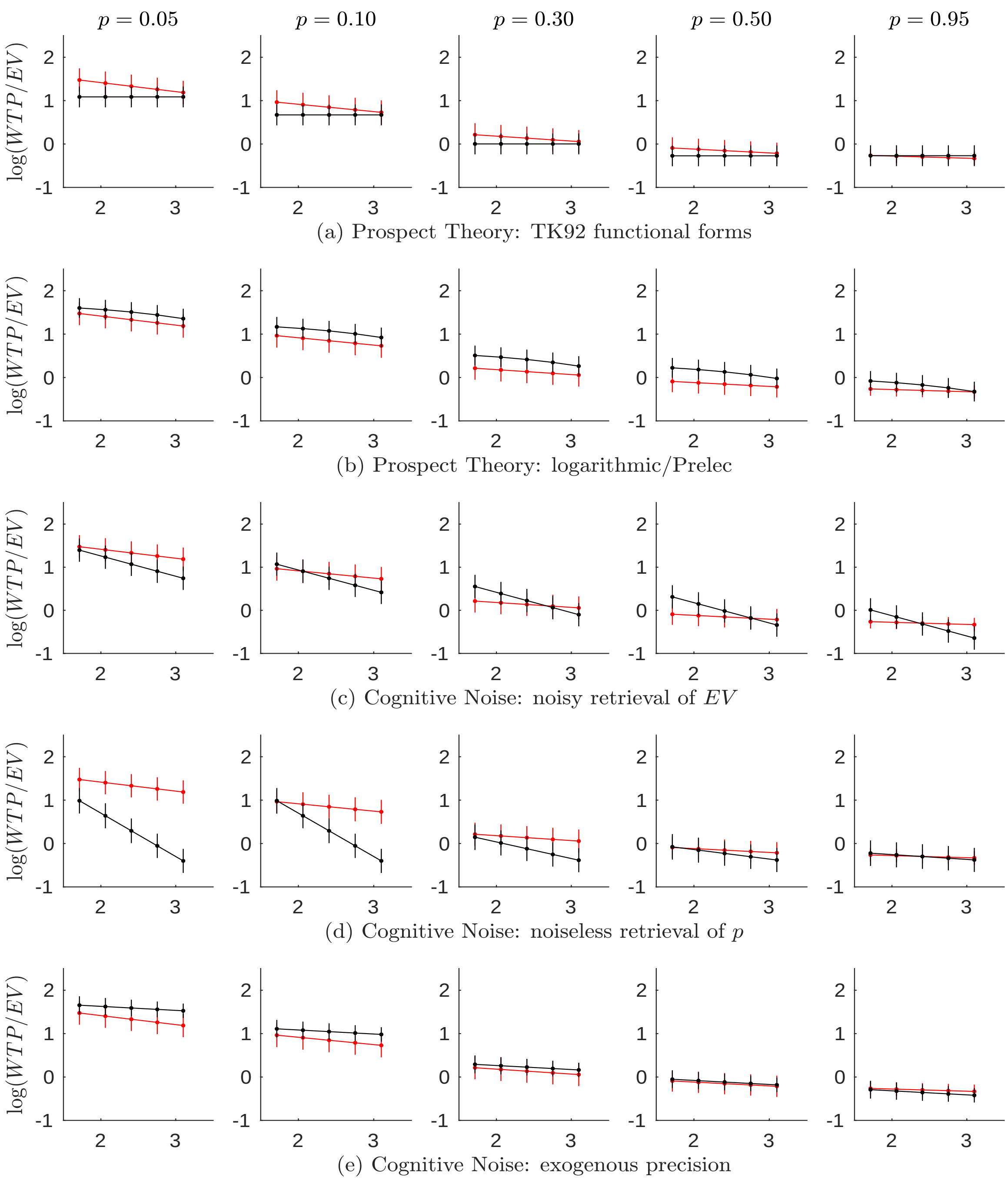

Figure 7: Predictions of alternative stochastic models of lottery valuation, using the same format as in Figures 5-6 (but only for selected values of p). Each row shows the predictions of the baseline model (in red) and one alternative model (in black).

is simply that the fit of our model is at least comparable to that of purely descriptive models that are commonly used to model data like ours. And rather than a competitor to prospect theory, our model is better understood as proposing a theoretical explanation for the nonlinear distortions of the objective data regarding payoffs and probabilities posited by PT. By deriving these distortions as features of an optimal decision rule when judgments must be based on a noisy internal representation rather the objective data, our analysis helps to explain why the kind of biases summarized by PT should be robust features of human decision making. It also offers the prospect of increased accuracy in empirical applications of PT, by providing insight into the circumstances under which the biases captured by PT should be most important (i.e., ones where one should expect greater noise in internal representations), and into the way in which the value function and weighting function of PT might be expected not to remain stable across environments.[76] Our comparison above of the average behavior of two groups of subjects who perform different versions of our experimental task provides a simple illustration.

4.2 Comparing Alternative Models of Cognitive Noise

We have focused thus far on global measures of fit for our complete model, taken as a package. Here we consider the contribution that particular features of our baseline model make to its empirical success. This allows us to address a question posed in the introduction: which kinds of cognitive noise are most important for explaining the variability of apparent risk attitudes?

Table 3 reports model comparison statistics for a variety of models, in each of which subjects' bidding rules are assumed to be optimally adapted so as to minimize the objective (2.7); but the models differ in their specification of cognitive noise. The top line recalls the log likelihood and BIC statistic for our baseline model, with three types of cognitive noise and endogenous imprecision in the representation of the monetary payoff. (These numbers are the same as those reported in Table 1 and again in Table 2.) The next line instead considers an asymmetric version of the model, in which the three noise parameters are allowed to be different in the case of lotteries involving losses rather than gains. While the log likelihood is necessarily slightly higher in the case, it is not enough higher to outweigh the penalty for the additional free parameters in the asymmetric case; the BIC is higher, implying a Bayes factor of 18 in favor of the symmetry assumption in our baseline model.[77]

The "exogenous precision" model instead assumes that ν_x is a fixed parameter for all lotteries, rather than varying with r_p as in the baseline model; the numerical value of ν_x is then a parameter to be estimated (instead of the cost function parameter A). This alternative, which requires stake-size effects to be of the same size for all p (since γ_x must be independent of r_p), reduces the likelihood of the data modestly, but does not dramatically worsen the fit of the model. (Row (e) of Figure 7 illustrates the extent to which the predictions of this model remain similar to those of the baseline model, even though it fails to capture the fact that β_p is more negative for the low values of p.)

[76] See section 4.3 for further discussion.

[77] The remaining theoretical models in the table continue to assume the same noise parameters in the case of both gains and losses.

model	#params	LL	BIC	K
baseline model	3	-1602.5	3236.3	1
asymmetric	6	-1598.1	3242.1	18
exogenous precision	3	-1604.7	3240.7	9.0
no payoff noise	2	-1608.7	3242.4	21
no probability noise	2	-1990.0	4005.0	8.3×10^{166}
no response noise	2	-1646.1	3317.2	3.7×10^{17}
noisy retrieval of EV	2	-1966.4	3957.9	4.9×10^{156}

Table 3: Model comparison statistics for alternative specifications of the cognitive noise model.

The next set of alternatives each shut off one of the types of cognitive noise in the baseline model. The model with "no payoff noise" assumes that the value of X is encoded and retrieved with perfect precision; this corresponds to a limiting case of the exogenous noise model in which $\nu_x = 0$ (or of the baseline model in which $A = 0$).[78] The model with "no probability noise" instead assumes that the value of p is encoded and retrieved with perfect precision (i.e., that $\nu_z = 0$), but still allows for noisy coding of the monetary payoff (with ν_x optimally determined as a function of p) as well as response noise. And finally, the model with "no response noise" assumes that ν_c, so that the DM's bid is optimally chosen as a function of the internal representation $\mathbf{r}$; the noisy internal representations of p and $|X|$ are specified as in the baseline model.

We find that any of these more restrictive models fits noticeably worse than the model with all three kinds of cognitive noise. Among the three, however, the assumption of noisy coding and retrieval of the payoff values is least crucial for the model's fit; while assuming perfect retrieval of the value of X reduces the likelihood of the data by more than 6 log points, once one penalizes the more flexible model for its additional free parameter, the Bayes factor[79] in favor of the baseline model relative to this alternative is only around 21. Eliminating response noise (while keeping both kinds of noise in the internal representation) reduces the likelihood (and so raises the BIC) a good deal more. Most important of all is the noisy coding of probabilities: assuming that $\nu_z = 0$ lowers the likelihood of the data to such an extent that even allowing for both noisy coding of payoffs (with a precision that depends on the value of p) and response noise, the Bayes factor in favor of the full model against this alternative is larger than 10^{166}.

The problem with this model is illustrated in row (d) of Figure 7: the model predicts that when $|X| = \bar{X}$, the prior mean value for $|X|$, the conditional mean $\mathrm{E}[\log(WTP/EV)\,|p, \bar{X}]$ should be the same for all p — the value of p affects only the slope of the line passing through that point.[80] This means that for values of X near its prior mean, such a model

[78] We discuss lottery valuation with no payoff noise in the Appendix, section B.2. But in the simple discussion there, we also abstract from response noise; the model considered on the third line of Table 3 allows for both noisy coding of probabilities and response noise.

[79] Note that in Table 3, the value of K indicates the factor by which each model is *inferior* to the baseline model. This convention is opposite to the one used in Tables 1 and 4, where K indicates the factor by which each model is *superior* to the reference model (e.g., the unrestricted model in Table 1).

[80] See the Appendix, section C.4, for the derivation of this prediction. The point at which the prediction

cannot account for the effects of changes in p on subjects' apparent risk attitude — the "fourfold pattern" of Kahneman and Tversky. Since many of the monetary payoffs used in the experiment are smaller than $\bar{X}$,[81] the pattern is to some extent captured by exaggerating the degree to which β_p is negative for small p; hence the best-fitting model of this type exaggerates the stake-size effects for small p, as shown in the figure.

The final line of the Table 3 considers an alternative model in which it is assumed the process of expected value computation has access to the precise values of p and X specified by the experimenter, but that the result of this computation is a noisy reading of the lottery's true EV (i.e., the quantity pX). While the sign of the EV is assumed to be recognized without error, the DM's bid is assumed to be based on a noisy semantic representation of the magnitude $|EV|$, drawn from a distribution

$$r_{ev} \sim N(\log |EV|, \nu_{ev}^2),$$

by analogy with our model (2.1) of the noisy internal representation of monetary payoffs. The DM's bid is drawn from a distribution (2.3), where the function $f(r_{ev})$ is optimally chosen to minimize the objective (2.7) as in our other models. This is a model with only two kinds of cognitive noise — noise in correctly retrieving the correct EV, and response noise given the DM's subjective sense of the lottery's value — and correspondingly two free parameters (ν_{ev} and ν_c). The predictions of this model, illustrated in row (c) of Figure 7, fit the bidding of the average subject considerably worse than those of our baseline model. The model implies the same stake-size effects for all p; but worse, it ties the strength of stake-size effects to the size of the effect of reductions in p on the relative risk premium, resulting in both an exaggeration of the predicted stake-size effects and an underestimation of the predicted effects of changes in p. Note that we are able to strongly reject this model only because our dataset includes separate variation in both p and $|X|$; we would not be able to discriminate between our baseline model and a model of noisy EV retrieval if we considered only a set of lotteries in which the payoff size varies with no variation in p (as in Khaw *et al.*, 2021) or a set of lotteries in which the probability varies but with no variation in the monetary payoff (as in Enke and Graeber, 2023).

4.3 Dependence of Parameters on the Number of Trials

Above we have fit the parameters of our cognitive noise models to the data moments for an "average subject," but we have seen from Figure 4 that there is heterogeneity in subjects' bidding behavior. Such heterogeneity is not necessarily inconsistent with the hypothesis of an optimal bidding rule, however, if we suppose that the cognitive noise parameters need not be identical for all subjects. As an illustration of this, we estimate the parameters of our baseline cognitive noise model separately for two different "average subjects," one based on the 13 subjects who each evaluated 400 lotteries, and the other based on the 15 subjects who each evaluated 640 lotteries. (We have already shown in Figure 4 that there is a systematic difference in the bidding by subjects in these two groups.)

is insensitive to the value of p corresponds to a value of $\log |X|$ slightly greater than 3 on the horizontal axis of the panels in row (d) of Figure 7.

[81]The skewness of the log-normal prior distribution implies that the prior mean is larger than the median payoff magnitude in the experiment.

Parameter Estimates: Baseline Cognitive Noise Model					
data	A	ν_z	ν_c	LL	LL/N
400-trial avg. subject	0.000004	1.19	0.29	-1202.6	-3.015
640-trial avg. subject	0.0078	2.28	0.28	-1995.8	-3.171
both average subjects	0.0021	1.75	0.27	-3249.8	-3.161
single average subject	0.0017	1.60	0.24	-1602.5	-3.068

Alternative Models of Both Average Subjects			
model	LL	BIC	K
common parameters	-3249.8	6534.2	1
separate parameters	-3198.4	6459.0	2.1×10^{16}

Table 4: Alternative estimates of the cognitive noise parameters for the baseline model, depending which average subjects' bidding behavior the model is required to explain. The upper part of the table presents the parameter estimates and a measure of the model's ability to fit each set of behavioral moments. The bottom part of the table compares two alternative uses of the model to explain the joint behavior of the 400-trial and 640-trial average subjects: one in which separate parameters are estimated for each average subject, and another in which the parameters are constrained to be the same for both.

The upper part of Table 4 shows how the estimated cognitive noise parameters differ across four possible versions of our model: a model fit only to the data of the 400-trial "average subject"; a model fit only to the data of the 640-trial "average subject"; a model fit to the data moments of the two "average subjects" together, but with a single set of parameters required to explain the behavior of both; and a model fit to the data moments of a single overall "average subject" (the baseline model in Table 3). For each of these estimation exercises, the maximized LL of the data moments is reported. The final entry in each row reports the value of LL divided by N, the number of observations in that dataset. This allows us a measure of the degree to which the optimizing model is able to fit the average subjects' behavior that is comparable across the different cases, despite the differing number of observations that are used to compute LL in the different cases.

We observe that the parameter values that best fit the behavior of the 640-trial average subject are fairly different from those that best fit the behavior of the 400-trial average subject: the 640-trial average subject has a much larger cost of precision in magnitude representation (and hence less precise representations of the monetary payoffs), and noisier internal representations of the probabilities as well, though the degree of response noise is similar for both. Moreover, the best-fitting parameters for either of the two groups are fairly different from those estimated when we require a single set of parameters to fit both average subjects (third line of the table), or when we fit the model to an average subject that pools the data from both groups of subjects (the bottom line). The cognitive noise model fits better (in the sense of achieving a high value of LL/N) to moments of the 400-trial average subject than to the data moments of the average subject when all 28 subjects were considered as a single group.

The bottom part of Table 4 demonstrates the value of allowing for heterogeneity in the parameters of the two groups through a formal model comparison. We consider two possible quantitative models of the 260 data moments consisting of the 100 data moments of the 400-trial average subject and the 160 moments of the 640-trial average subject. In one model (the "separate parameters" model), we fit the model separately to the moments of each of the two average subjects; the best-fitting parameter values for each subject are the ones shown on the first two lines of the upper part of the table. The LL for this model is just the sum of the LLs shown on those two lines. In the other model (the "common parameters" model), we instead require the values of the parameters to be the same for both average subjects; the best-fitting parameter values for this exercise are shown on the third line of the upper part of the table. The LL for this model is also taken from the third line in the upper part of the table. Since the two models involve different numbers of free parameters, we compare their degree of fit using the BIC rather than the LL alone. Because the "common parameters" model is more parsimonious, the difference in the BICs of the two models is not as great as twice the difference in their LLs. Nonetheless, the "separate parameters" model fits the data better, even using the BICs as the basis for our judgment. The implied Bayes factor in favor of the "separate parameters" model is greater than 10^{16}.

Thus we can improve the fit of the model, relative to what is indicated by the fits shown in Figures 5 and 6, by allowing separate parameters for the two groups of subjects. Moreover, the nature of the difference in the parameter values for these two groups of subjects makes a certain amount of sense, given the greater mental fatigue or loss of concentration that one might expect in the case of the subjects who were required to complete a substantially longer series of trials. Requiring more trials appears to reduce the precision of the internal representation of both the probabilities and the monetary payoffs, but with a more dramatic effect on the representation of the monetary payoffs. Heterogeneity of this kind in our dataset is quite consistent with our interpretation of departures from risk-neutral bidding as an adaptation to cognitive noise. It is instead less obvious, in the context of a purely descriptive model such as prospect theory, why the parameters of the model should vary systematically between the subjects selected to face different numbers of and distributions of lotteries.

5 Discussion

We have shown that a model in which subjects' bids are hypothesized to be optimal — in the sense of maximizing the DM's expected financial wealth, rather than any objective that involves true preferences with regard to risk, and without introducing any free parameters representing such DM preferences — can account well for both the systematic biases and the degree of trial-to-trial variability in our subjects' data, once we introduce the hypothesis of unavoidable cognitive noise in their decision process. The model can simultaneously account for the fourfold pattern of risk attitudes predicted by prospect theory (Tversky and Kahneman, 1992), relating to the effects of varying payoff probabilities and the sign of the payoffs, and the alternative fourfold pattern of Hershey and Schoemaker (1980) and Scholten and Read (2014), relating to the effects of varying payoff magnitudes and the sign of the payoffs. The effects on the sign of the relative risk premium of varying the terms of a simple lottery along each of these three dimensions can be explained by a single theory, which attributes departures from risk neutrality in either direction to the way in which bids should be shaded in order to take account of cognitive noise. Thus stake-size effects are shown not only to be consistent with the classic effects emphasized by prospect theory, but even to have the same underlying explanation.

We have modeled a decision process with cognitive noise by supposing that the process computes a quantity $f(\boldsymbol{r})$ on the basis of a noisy internal representation $\boldsymbol{r}$ of the data defining the decision problem; this quantity $f(\boldsymbol{r})$, which may be thought of as an imprecise subjective representation of the value of the lottery, must then be matched (in a way that involves random error in recognizing the correspondence between the subjective sense of value and a particular monetary amount) with a bid C that expresses that valuation. Noise enters through the way in which $\boldsymbol{r}$ is determined, conditional on the true lottery characteristics (p, X), and through the way in which the bid C is determined given the value of $f(\boldsymbol{r})$; the function $f(\boldsymbol{r})$ is assumed to be computed precisely, and to be optimal given the statistics of the DM's situation. At least within the context of our model, we find that both kinds of noise (the encoding noise in the determination of $\boldsymbol{r}$ given the true data, and the response noise in the determination of the bid given $f(\boldsymbol{r})$) are needed to account for our data; and we find that a model with separate noise in the encoding of the quantities p and $|X|$ individually fits much better than a model with noise only in reporting the expected value computed from the data (p, X).

The structure of our theoretical model is similar to that of a Bayesian model of bias in the estimation of sensory magnitudes (e.g., Petzschner *et al.*, 2015). But this does not mean that the cognitive noise assumed in our theory is *perceptual* error, like the random signals from sensory receptors in the case of physical properties of the environment. Our assumption that $\boldsymbol{r}$ is random conditional on the true data (p, X) mean that there is error in the DM's perception of what the experimenter has told them; presumably this is minimal in the case of an experiment where both payoffs and probabilities are described to the subject using number symbols.[82] Rather, it simply means that even though the DM understands the numbers that they have been told correctly, they do not use this knowledge in an explicit arithmetic calculation, but instead form an intuitive judgment about the value of the presented options, drawing upon an imprecise "semantic representation" of the numbers that they have been told.[83] The semantic representation is drawn upon because it is easier to combine several pieces of information to produce an intuitive judgment of overall value using this system; answering this type of question using the DM's explicit knowledge of the precise numbers (in another part of their brain) would be possible only through a more laborious process, accessing learned arithmetic facts and abstract ideas about the connection between lottery valuation and mathematical calculation. Yet the same person may well answer simpler questions ("what did I say you will win if the lottery pays off?") using their ability to recognize the exact meaning of number words, indicating that they did correctly see the information on the screen.

There are a variety of reasons to believe that many aspects of the apparent risk attitudes measured in laboratory experiments reflect an adaptation to the presence of cognitive noise, rather than considered preferences with regard to risk; these go beyond our demonstration

[82] The experiment of Barretto-García *et al.* (2023) described below, in which payoffs are indicated by an array of coins, instead introduces an element of genuine perceptual error.

[83] On the co-existence of an imprecise semantic representation of numbers alongside precise number information, in different brain areas, see Dehaene (2011).

here that an optimizing model can fit experimental data on lottery valuations at least as well as quantitative versions of prospect theory. Notably, measured risk attitudes are correlated with a variety of quantitative indicators of cognitive imprecision. One such measure is the degree to which the same subject is observed to give variable responses on repeated presentations of the same data. Khaw *et al.* (2021) show that across their subjects, there is a significant positive correlation between an estimated subject-level index of risk aversion in binary choices between lotteries and a subject-level index of the stochasticity of choice, exactly as the cognitive noise model would predict if the difference in subjects' choice behavior is due mainly to subject-level differences in the amount of cognitive noise; Barretto-García *et al.* (2023) replicate this finding using a similar experimental design.[84] Oprea (2023) considers lottery valuation rather than binary choice, and a broader range of types of lotteries, and also finds that prospect-theoretic deviations from valuation of a lottery in accordance with its EV are greater for those subjects whose responses are more variable from trial to trial.

Another possible indicator of the degree of cognitive noise, that does not require repeated presentations of the same decision problem, is a subject's reported degree of uncertainty about the correct response to give on a single trial. The theoretical model that we have presented does not necessarily require that the DM be aware of the degree to which their internal representations are noisy.[85] On the other hand, we might easily assume that the DM has access both to summary statistics r_p and r_x (the only information needed to choose their bid on that trial) and to a sense of how noisy these readings are, without knowing the exact distribution from which they have been drawn (which would reveal the true values of p and X). For example, suppose that the internal representation of $|X|$ is actually given by the sample path of a Brownian motion with a drift proportional to $\log |X|$ and instantaneous variance independent of $|X|$;[86] then the posterior distribution for $|X|$ depends only a scalar summary statistic r_x (the cumulative distance traveled by the Brownian motion), but the instantaneous variance (and hence the degree of cognitive noise) could also be determined from this internal representation. In such a case, one might well suppose that the DM should be able to assess their degree of uncertainty. Enke and Graeber (2023) show that subjects in a lottery valuation experiment do indeed have some degree of metacognitive awareness of the imprecision of their valuations, and show furthermore that prospect-theoretic deviations from risk-neutral valuation (in each of the four quadrants of the "fourfold pattern") are stronger for those subjects who express greater uncertainty;[87] Oprea (2023) finds the same

[84] The lotteries used in these studies all involve a probability $p > 0.5$ of a positive gain, so that prospect theory would predict risk-averse valuations; and this is the direction of deviation that is observed in all of the subjects who differ notably from risk-neutral choice. The subjects whose choices are most predictable are also the ones for whom the "certainty equivalent" of a risky lottery — defined as the certain amount for which the subject is equally likely to choose the lottery and the certain amount — is closest to its EV.

[85] We assume that the DM has learned to calibrate their intuitive responses to those representations in a way that is optimal, given the degree to which the representations are noisy, but this need not involve conscious knowledge of how their intuitive judgments are formed, or why a response shaded in that way is optimal.

[86] This hypothesis provides a possible justification for the cost function (2.6), as discussed in Appendix section A.

[87] See our Appendix, section B.2, for a model of how the results of Enke and Graeber (2023) can be explained by a model of optimal adaptation to noise in the internal representation of the relative probability

using a similar method.

Oprea (2023) also finds that subjects' degree of departure from risk-neutral valuation is positively correlated with their self-reported degree of inattention to the data on payoffs and probabilities that they have been shown, and their self-assessment of the degree to which they have "guessed" rather than making a "precise (exact) decision." It is also negatively correlated with the average time that a subject takes to respond, which can be taken as an indicator of the amount of cognitive effort expended on ensuring accuracy (Rubinstein, 2013);[88] and positively correlated with the number of errors that a subject makes on a cognitive reflection test, which can also be taken as a measure of cognitive imprecision. Note that under an interpretation of prospect-theoretic biases as simply reflecting non-standard preferences, there would be no reason to expect any of these correlations with indicators of cognitive imprecision.

The cognitive noise posited in our model (and indicated by the findings just summarized) may be specific to the way in which people assess the value of gambles. But we have based our quantitative specifications (2.1) and (2.2) on a proposed analogy with the measured imprecision of number processing in other contexts, especially errors in estimation and discrimination when information about numbers or proportions is presented through visual displays; and the success of our model suggests that the noise may relate to number processing more generally. Barretto-García *et al.* (2023) provide evidence that this may be the case, by having the same subjects choose between lotteries in which the monetary payoffs are specified by number symbols; between lotteries in which the payoffs are instead presented visually (by a irregular array of euro coins); and choose the larger of two numbers shown by visual arrays (a pure numerosity discrimination task). They find similar patterns of stochastic choice (risk-neutrality and scale-invariance) regardless of the way in which the payoffs are presented, except that there is greater apparent noise in the internal representation of the payoffs (and correspondingly greater apparent risk aversion) when the payoffs are presented visually. Moreover, the imprecision of individual subjects' numerosity discrimination (which can be taken to indicate the degree of noise in their internal semantic representation of numbers) is correlated both with the noisiness of their choices between lotteries and with their degree of apparent risk aversion.[89] These authors also seek to measure the degree of precision of subjects' neurocognitive number representations more directly by having them look at numbers while in an fMRI scanner, and find a significant correlation between this measure of imprecision in number representation and both the degree of noise in subjects'

of the two outcomes. The association that they find between subjective uncertainty and the size of the departure from risk-neutral valuation is the one that our model would predict, if subjects differ in the value of ν_z and have some degree of access to the size of their personal noise parameter.

[88]In some cases, however, a longer average response time is taken to indicate a more difficult decision, and is expected to be associated with more random choices (e.g., Alós-Ferrer *et al.*, 2021). The correlation found by Oprea (2023) can be interpreted as indicating that the main reason for differences in subjects' average response times is subject-level differences in the amount of concern for precision, rather than differences in the intrinsic difficulty of the problems faced by different subjects.

[89]Frydman and Jin (2022) similarly find a significant correlation between the imprecision of subjects' number comparisons and the same subjects' degree of departure from risk neutrality in lottery choice. In the experiment of Frydman and Jin, the number-comparison task uses symbolically presented numbers rather than visual arrays, but requires rapid answers; it seems that in such a case, an imprecise semantic representation of the numbers is drawn upon because it allows faster answers (Dehaene *et al.*, 1990).

lottery choices and their degree of departure from risk-neutrality.

Another telling source of evidence that apparent risk preferences may reflect cognitive imprecision is the observation that they are unstable, changing with experience and feedback. According to the "discovered preference hypothesis" of Plott (1996), subjects in laboratory experiments can only be expected to consistently choose in accordance with their actual, considered preferences after sufficient experience with the form of task used in the experiment, involving feedback as to the consequences of their choices; and indeed, a number of authors have found that the apparent risk preferences that are expressed in relatively novel choice problems are unstable in this way (the so-called "description-experience gap": Hertwig and Erev, 2009). Perhaps the most celebrated finding of this kind is the observation that the tendency of subjects to overweight small-probability extreme outcomes, as predicted by prospect theory, occurs when lotteries involving such extreme outcomes are described to subjects (as in the experiment of Tversky and Kahneman, 1992), but disappears when subjects learn about the outcomes from repeated choices with feedback as to the consequences of choosing the lottery or not on each occasion (Hertwig *et al.*, 2004).[90] Van de Kuilen and Wakker (2006) find that the Allais Paradox disappears with sufficient experience and feedback, and Van de Kuilen (2009) finds more generally that evidence of nonlinear probability weighting disappears. Ert and Haruvy (2017) and Charness *et al.* (2023) further find that sufficient experience and feedback greatly reduce measured departures from risk-neutrality.

These findings are consistent with the theory proposed here, if we suppose that experience with a particular format of decision problem and feedback regarding outcomes from one's choices reduces the degree of noise in the DM's internal representation of the lottery, and that people have sufficient awareness of the increasing precision of their understanding of the decision problems to adapt their decision rule in a way that is optimal for a lower-noise situation (resulting in more nearly risk-neutral choices). This proposal would be consistent with Plott's discovered preference hypothesis, but extends it: whereas Plott is concerned simply to assert that rational choice theory should be descriptive over a sufficiently restricted domain (sufficient experience, sufficient feedback, etc.), we also offer a theory of the kind of departures from normatively correct choices that should be observed when there is little experience with the necessary kind of feedback (as in our experiment).

The studies just mentioned show that it is possible to reduce the imprecision of internal representations through proper experimental design; but it is equally possible to design experimental treatments that should predictably *increase* cognitive imprecision, and this possibility is of particular interest as a test of our theory. For example, Enke and Graeber (2023) increase the complexity of way in which information about the probability p is presented to their subjects (requiring the subjects to correctly reduce a compound lottery to recognize the implied probability of a non-zero payoff, rather than telling them the number directly), and show that this increases the strength of prospect-theoretic biases in their subjects' lottery valuations — in exactly the way that our model would imply in the case of

[90]Hertwig *et al.* (2004) find that when subjects must learn about the distribution of possible outcomes purely from experience, the typical result is *under-weighting* of low-probability extreme outcomes, rather than perfectly risk-neutral choice. This bias can be explained, however, as reflecting the fact that in a small sample of experience the low-probability extreme outcome will often happen not to have been observed. The bias may well reflect incomplete learning — and subjects' choices may well be optimal, conditional on the imprecise representation of the lottery characteristics that is afforded by a finite sample of experience.

an increase in the imprecision of the internal representation of probabilities.[91] Subjects have similarly been shown to depart farther from risk-neutral choice as a result of increased time pressure (Choi *et al.,* 2022; Kirchler *et al.,* 2017, Young *et al.,* 2012), increased cognitive load (Benjamin *et al.,* 2013; Deck and Jahedi, 2015; Gerhardt *et al.,* 2016), or acute stress (Porcelli and Delgado, 2009). These are all plausibly conditions that can be expected to reduce the precision of mental processing, perhaps in ways that can be captured by an increase in the noise parameters in our model;[92] if so, the model implies that we should expect larger departures from risk-neutrality in exactly the directions that are observed.

Finally, there is evidence that the valuation biases observed in experiments like ours also occur in problems that have nothing to do with risk. Oprea (2023) has subjects answer questions both about their preferences over lotteries, and about "deterministic mirrors" of the same choice problems: in addition to eliciting the amount C such that the subject would be indifferent between receiving C for sure and receiving the outcome of a lottery promising X with probability p (the "certainty equivalent" of the lottery), he elicits the amount C such that the subject would be indifferent between receiving C and receiving a fraction p of the larger amount X (the "simplicity equivalent" of the more complex offer). In both cases, the subject would maximize their expected financial reward by making choices consistent with an indifference value of pX; but the second kind of problem involves no risk. Oprea shows that elicited indifference values differ from pX in exactly the same way in both kinds of problems (the "fourfold pattern of risk attitudes" appears also in the deterministic problem), and even to a quantitatively similar extent. Moreover, the degree to which individual subjects differ from reward-maximizing choice is highly correlated across the two tasks; and individual differences in the degree of departure from reward-maximizing choice in *both* tasks are correlated with indices of cognitive imprecision (variability of responses on repeated trials, reported uncertainty about the response, self-reported inattention to the problem data, faster responses, etc.) in the way discussed above.

Vieider (2023) finds similarly parallel biases in a comparison of binary choices between lotteries and corresponding "deterministic mirror" problems. And Enke and Shubatt (2023) show that differences in the degree to which people make errors in judging the expected value of a lottery (when rewarded for their answer to this arithmetic problem, and not required to accept any risk) can be used to predict the degree to which they make choices inconsistent with EV maximization when asked to choose between lotteries.[93] This close parallelism between errors in arithmetic problems that involve no risk and apparent "risk attitudes" in choices involving lotteries is exactly what our model of lottery valuation would imply, if

[91] See the Appendix, section B.2. In a different type of decision problem, Charles *et al.* (forthcoming) also find that the strength of a cognitive bias (imperfect pass-through from beliefs to actions) can be changed by manipulating the complexity of the information upon which subjects' beliefs should be based.

[92] Our finding in this paper that both prospect-theoretic valuation biases and stake-size effects are stronger for the subjects who completed a larger number of trials (recall Figure 4 and the results in section 4.3) suggests that fatigue may have a similar effect.

[93] Enke and Shubatt (2023) further use subjects' rates of error in different EV calculations to estimate a machine-learning model of the "complexity" of a given EV calculation, and show that this measure can also be used out of sample to predict the degree of noise in lottery choice. The measure of complexity that they propose, however, is purely empirical. The model of noisy coding and optimal decoding presented here can be viewed as providing a theory of what determines the "complexity" of a lottery, in the case of simple lotteries of the kind used in our experiment.

one assumes that the quantities p and X defining the "deterministic mirror" of a lottery are encoded and decoded in a similar way as the quantities defining a lottery, given that we posit a decision rule that minimizes the mean squared error (2.4).

We have noted in the introduction that rather than offering a theory that is completely different from prospect theory, our noisy coding model can be viewed as providing an interpretation of the source of the nonlinear transducers posited in prospect theory, and indeed one that is in line with (though providing a more detailed account than) the discussion by Kahneman and Tversky themselves of the way that the value function and probability weighting function reflect a sort of psychophysical diminishing sensitivity. It might be asked then whether the noisy coding model adds anything to quantitative versions of prospect theory. We believe that our account deepens understanding of the implications of prospect theory in two important respects.

First, interpreting the nonlinear distortions posited by prospect theory as consequences of optimal adaptation to cognitive noise helps to explain the observed instability of prospect-theoretic parameters across settings, as sketched above. While a complete account of the way in which the degree of cognitive noise should vary across settings, and of the rate at which decision rules should be expected to adapt to changes in the degree of cognitive noise, remains to be elaborated in future work, this interpretation offers the prospect of a more general theory that can explain when and to what degree one should expect prospect-theoretic parameters measured under one condition to predict behavior under other conditions. In this way, one can hope to use prospect theory more accurately as a predictive tool, somewhat in the spirit of the "Lucas critique" (Lucas, 1976; Sargent, 1987) of the naive use of econometric relationships for policy evaluation.

And second, our interpretation cautions against the use of estimated prospect-theoretic "preferences" as a basis for welfare evaluation. To the extent that the parameters of prospect-theoretic transducers are found to change with experience and feedback, as discussed above, one should doubt that the apparent preferences estimated in situations where there has been little experience or feedback really represent considered preferences. Our analysis provides further grounds for such skepticism by showing that, even when behavior is only observed in novel situations where feedback is not given, there is good reason to regard the anomalous patterns of behavior summarized by prospect theory as reflecting errors due to cognitive noise. And our theory provides an alternative basis for welfare judgments, even if subjects' choices under ideal conditions of extensive experience and feedback cannot be observed. To the extent that one can show that peoples' choices appear to have been optimized to achieve a particular objective (here, the objective of maximization of expected financial wealth), one might plausibly regard that objective as reflecting their "true" preferences. This would still provide a basis for welfare judgments that is individualistic, in the sense that what is good for a person is inferred from what their observed behavior apparently aims to achieve, rather than reflecting what someone else wants for them (Woodford, 2018).

These cautions about the uses of prospect theory in policy design hardly imply that the departures from normative behavior documented by authors like Kahneman and Tversky are of no importance for policy analysis. The claim that people would behave in closer conformity with normative decision theory with sufficient experience and feedback should no more justify assuming that one can always assume such behavior when designing policies than the assertion that wages and prices will adjust "in the long run" so as to make real quantities independent of monetary variables would justify indifference to the extent to which different monetary policies stabilize aggregate nominal spending. We expect that the design of economic policies can be improved by taking into account the ways in which people are prone to misunderstand the circumstances under which they act. But the realization of this promise will require further progress in understanding the nature of cognitive noise and the way in which people adapt their behavior to deal with it.

ONLINE APPENDIX

Khaw, Li, and Woodford, "Cognitive Imprecision and Stake-Dependent Risk Attitudes"

A The Cognitive Cost of Precision in the Representation of Monetary Amounts

Our baseline model assumes that is possible to vary the precision with which monetary amounts are represented, subject to a cost of the form (2.6). This specification of the cost function has a simple interpretation. Suppose that the magnitude $|X|$ is internally represented by a random quantity that evolves according to a Brownian motion, with a drift equal to $\log|X|$ and an instantaneous variance $\sigma^2 > 0$ that is independent of $|X|$. (It suffices for our argument that the drift be an affine function of $\log|X|$, but the calculations are simplified by assuming that the drift is simply equal to $\log|X|$. The assumption that $y_0 = 0$ is also purely to simplify the algebra.) This process y_t is allowed to evolve for some length of time $\tau > 0$, starting from an initial value $y_0 = 0$; the final value y_τ constitutes the internal representation. Diffusion processes of this kind are often used to model the randomness in sensory perception and memory retrieval.[94]

Equivalently, we may treat the value $r_x \equiv y_\tau/\tau$ as the internal representation, as this variable contains the same information as y_τ. Under this assumption, the internal representation has the distribution specified in (2.1), where $\nu_x^2 = \sigma^2/\tau$. Note that the precision of such a representation can be varied by varying τ, the length of time for which the process y_t is allowed to evolve. Moreover, successive increments of the Brownian motion are independent random variables (with a common distribution that depends on the magnitude $|X|$); these can be thought of as repeated noisy "readings" of the value of $|X|$.[95] If we suppose that each repeated "reading" has a separate (and identical) psychic cost, then the total cost should be linear in τ (and so proportional to the total number of independent "readings"). This implies a cost of precision of the form (2.6).

B The Log-Odds Model of Noisy Coding of Probabilities

Here we provide further justification for our interest in the model (2.2) for the noisy internal representation of the relative probability of the two possible outcomes of a lottery.

[94] See Gold and Heekeren (2014) for a review. Heng *et al.* (2023) use a process of this kind to model the internal representation of positive numbers presented as arrays of dots, and show that the assumption of precision increasing linearly with time fits well the way that the distribution of errors in numerosity estimation varies with viewing time.

[95] Gold and Heekeren (2014) discuss the neural mechanisms that could implement such a process.

B.1 Consistency with the Logarithmic Model of Encoding of Positive Magnitudes

We first note that our model (2.2) of the imprecise representation of probability information is closely related to the way in which we model the imprecise representation of numbers, in our discussion of the internal representation of the monetary payoffs offered by a lottery.

Suppose that the relative probability of the two possible outcomes is displayed to a subject by the relative size of two magnitudes, X_1 and X_2, proportional to the probabilities of the two outcomes. (In the case of our experiment, X_1 and X_2 could be the lengths of the two bars corresponding to the probabilities of the two outcomes, as shown in Figure 1.) And suppose that each of these magnitudes is independently encoded by a noisy internal representation, where

$$r_j \sim N(\log X_j, \nu_p^2), \qquad j = 1, 2,$$

as specified for the monetary amounts in (2.1). (Note that this would also be a common model of imprecision in visual perception of length.) Finally, suppose that judgments about the relative probability of the two outcomes are based purely on the *difference* between these two internal representations, $r_p \equiv r_1 - r_2$. In this case, the conditional distribution of the internal representation r_p of the relative odds will be of the form (2.2), where p in this expression means the probability of outcome 1, and $\nu_z^2 = 2\nu_p^2$. Note, however, that our conclusions in our baseline of model of lottery valuation depend only on assuming (2.2), and not on this particular interpretation of how the internal representation of the relative odds may be constructed.

Another reason for proposing (2.2) as a model of imprecision in the internal representation of probability information is the usefulness of a model of this kind in accounting for measured imprecision of people's judgments about probabilities, relative frequencies, and proportions, when these are presented visually or through a sample of instances (rather than with number symbols as in our experiment). The idea is parallel to our hypothesis about the encoding of numerical magnitudes: that the imprecision in the internal representation of numbers is the same when numbers are presented symbolically (as in our experiment) as in the better-studied case of judgments about numbers presented visually (numbers represented by the length of a bar, or the number of items in an array).

For example, Eckert *et al.* (2018) present evidence for a similar model of the discriminability of different relative frequencies of occurrence of two possible outcomes. They experimentally test the ability of both humans and chimpanzees to distinguish between two urns, one with a ratio $a_1 : a_2$ of outcomes of type 1 rather than of type 2, and another with a ratio $b_1 : b_2$ of the two possible outcomes, and find that the probability of correctly recognizing which offers the higher probability of outcome 1 is an increasing function of the "ratio of ratios" $(a_1/a_2)/(b_1/b_2)$. Equivalently, it is an increasing function of the difference in the log odds associated with the two urns,

$$\log \frac{a_1}{a_2} - \log \frac{b_1}{b_2}.$$

This is exactly the prediction of our model, if each relative frequency $p_i : (1 - p_i)$ is encoded by a noisy internal representation r_{pi} drawn from the distribution (2.2), and the DM's judgment is based on the relative size of r_{p1} and r_{p2}. Note also that Eckert *et al.*

conclude from their findings "that intuitive statistical reasoning relies on the same cognitive mechanism that is used for comparing absolute quantities, namely the analogue magnitude system." Here by "the analogue magnitude system" they mean a system of imprecise semantic representation of natural numbers, with the property that "discriminability of two sets varies as a function of the ratio of the set sizes to be compared, independent of their absolute numerosity,"[96] as would be implied if the two numerical magnitudes are encoded logarithmically as specified in (2.1).

B.2 An Optimizing Model of Bias in the Estimation of Probabilities or Relative Frequencies

Additional (though slightly less direct) evidence in favor of a model of noisy internal representation of probabilities like (2.2) comes from studies in which subjects must produce an *estimate* of some probability or proportion, rather judging which of two probabilities is greater. While studies of biases in estimation (as opposed to discriminability) provide less direct evidence about the precision of the internal representations on which the estimates are based, they are arguably of more direct relevance to the cognitive task in our experiment (i.e., producing an estimate of the value of a lottery).

Studies of bias in the estimation of probabilities, relative frequencies, and proportions find that people's estimates are typically most accurate for probabilities near 0 or 1, but much less accurate for intermediate probabilities (Hollands and Dyre, 2000; Zhang and Maloney, 2012). Moreover, at least in cases where there are only two possible outcomes, and the distribution of values for the probability of the first outcome is symmetric around 0.5, the degree of estimation error is typically found to be symmetric around 0.5, as the specification (2.2) together with a hypothesis of Bayesian decoding would imply. In fact, Zhang and Maloney (2012) review a wide range of previous experiments requiring subjects to judge the relative frequency with which two outcomes occur — either when presented simultaneously (say, a visual image containing many randomly arranged dots of two different colors) or in sequence (say, a succession of letters that are either of one type or the other). They show that characteristically, the median estimate $\bar{p}$ is a function of the true probability (or relative frequency) p of the form

$$\log \frac{\bar{p}}{1-\bar{p}} \;=\; \gamma \log \frac{p}{1-p} \;+\; (1-\gamma) \log \frac{p_0}{1-p_0}, \tag{B.1}$$

for some "anchor" or reference probability p_0 and an adjustment coefficient that in most cases satisfies $0 < \gamma < 1$. The reference probability p_0 is different in different experiments, but Zhang and Maloney note that it is typically close to the average of the true values p used in the experimental trials. Here we show how such a pattern of bias can result from optimal Bayesian decoding of a noisy internal representation of the kind specified by (2.2).[97]

Bayesian decoding of the noisy internal representation can only be defined relative to a prior distribution of true values of p for which the subject's decision rule has been optimized.

[96]See Eckert *et al.* (2018, abstract and p. 100).

[97]Zhang *et al.* (2020) propose a related model, and fit it to a variety of experimental datasets, though their model of the noisy coding of probability information is more complex, and their model of estimation on the basis of the noisy internal representation is not fully Bayesian.

A hypothesis that is convenient for such calculations (and that delivers a linear-in-log-odds relationship, at least approximately) is to assume a logit-normal prior,

$$z \sim N(\mu_z, \sigma_z^2), \tag{B.2}$$

where we introduce the notation $z \equiv \log(p/1-p)$ for the log odds. In the case of such a prior, the posterior distribution for the log odds, conditional on the representation r_p, will be a Gaussian distribution

$$z \sim N(\hat{\mu}_z(r_p), \hat{\sigma}_z^2), \tag{B.3}$$

where

$$\hat{\mu}_z(r_p) = \mu_z + \left(\frac{\sigma_z^2}{\sigma_z^2 + \nu_z^2}\right)(r_p - \mu_z), \qquad \hat{\sigma}_z^{-2} = \sigma_z^{-2} + \nu_z^{-2}. \tag{B.4}$$

It is not entirely clear what objective should be maximized by subjects' responses in the experiments reviewed by Zhang and Maloney (2012), since the experiments are typically not incentivized (and of course, one might assume in any event that there should be important "psychic" benefits from accuracy in addition to any monetary rewards). One simple hypothesis might be that the subject's estimate $\hat{p}$ is the one implied by the maximum a posteriori (MAP) estimate of the log odds of the event conditional on an internal representation r_p with statistics of the kind proposed above.[98] In this case, the model predicts an estimate

$$\hat{p}(r_p) = \frac{e^{\hat{z}(r_p)}}{1 + e^{\hat{z}(r_p)}}, \tag{B.5}$$

where the estimated log odds are given by $\hat{z}(r_p) = \hat{\mu}_z(r_p)$, the function defined in (B.4). We obtain the same prediction if instead we suppose that a subject computes an estimate of the log odds given by the posterior mean value of z, and then converts this into an implied estimate for p using (B.5).

Then since $\hat{\mu}_z(r_p)$ is a monotonic function, and the estimate for p specified in (B.5) is also a monotonic function of the estimate for z, the median estimate of p is predicted to be

$$\bar{p} = \hat{p}(z) = \frac{e^{\hat{\mu}_z(z)}}{1 + e^{\hat{\mu}_z(z)}}.$$

This implies that

$$\log \frac{\bar{p}}{1 - \bar{p}} = \hat{\mu}_z(z),$$

which is a relation of the form (B.1), in which

$$\gamma = \hat{\gamma} \equiv \frac{\sigma_z^2}{\sigma_z^2 + \nu_z^2}, \qquad \log \frac{p_0}{1 - p_0} = \mu_z. \tag{B.6}$$

The average estimated log odds would thus be an increasing function of the true log odds, with a slope less than one, implying a conservative bias. Moreover, the cross-over value is

[98] This hypothesis is discussed mainly because it allows us a simple closed-form solution. However, in at least some experimental studies of probability estimation, subjects report their probability estimate in terms of log odds; see Phillips and Edwards (1966). And Zhang and Maloney (2012) argue that there is reason to believe that the brain represents probabilities in terms of log odds, so that probability estimates can be understood as resulting from intuitive calculations in terms of log odds.

predicted to be the probability corresponding to log odds of $z = \mu_z$: the mean of the possible log odds under the prior. Hence this kind of Bayesian model provides a potential explanation for the results summarized in Zhang and Maloney (2012).

An alternative behavioral model would assume that subjects' estimates of p correspond to the posterior mean value of p (rather than the value of p implied by the posterior mean value of z); that is, that $\hat{p} = \mathrm{E}[p\,|r_p]$. In this case, we cannot give an explicit analytical solution for $\hat{p}(r_p)$, but Daunizeau (2017) offers a "semi-analytical" solution which he shows numerically is quite accurate over a wide range of parameter values. Using this result, the posterior expected value $\hat{p}$ can be approximated by the value such that

$$\log \frac{\hat{p}}{1-\hat{p}} \;=\; \alpha \hat{\mu}_z(r_p), \tag{B.7}$$

where

$$\alpha \;=\; [1 \,+\, a\hat{\sigma}_z^2]^{-1/2} \;<\; 1$$

and a is a constant equal to about 0.368. The median estimate of p should then satisfy

$$\log \frac{\bar{p}}{1-\bar{p}} \;=\; \alpha \hat{\mu}_z(z), \tag{B.8}$$

which is again a relation of the form (B.1), but now with

$$\gamma \;=\; \alpha\hat{\gamma}, \qquad \log \frac{p_0}{1-p_0} \;=\; \left(\frac{\alpha(1-\hat{\gamma})}{1-\alpha\hat{\gamma}}\right)\mu_z,$$

where $\hat{\gamma}$ is again defined as in (B.6).

Again we find (to an excellent degree of approximation) that the relationship between p and the median estimate $\bar{p}$ should be of the linear-in-log-odds form assumed in the regressions of Zhang and Maloney (2012). Again the average estimated log odds would thus be an increasing function of the true log odds, with a slope less than one, implying a conservative bias; and again the value of the log odds at which the cross-over from over-estimation to under-estimation should occur is an increasing function of μ_z (though deviating from even odds slightly less than does μ_z). The consistency of these results with the empirical evidence in Zhang and Maloney (2012) suggests that the model (2.2) of imprecise encoding of probability information is a realistic one.

Note that in an experiment like that of Enke and Graeber (2023), in which the magnitude $|X|$ is the same on all trials (with only p and the sign of X differing across trials), a model of bias in the estimation of probabilities directly implies a model of bias in lottery valuations. If $|X|$ is the same on all trials, and we assume a decision rule that is optimized for the distribution of lotteries actually encountered in the experiment, then there can be no posterior uncertainty about the value of $|X|$. Then if we ignore the issue of response error (analyzed in section C.2), the bidding rule that would maximize the DM's expected financial wealth will simply be

$$C \;=\; \mathrm{E}[p\,|r_p] \cdot X,$$

so that our model predicts

$$\log \frac{WTP}{EV} \;=\; \log \mathrm{E}[p\,|r_p] \;-\; \log p. \tag{B.9}$$

Thus the relative risk premium implied by subjects' bids should (according to our model) be purely a function of the bias in the optimal Bayesian estimate of p conditional on the noisy internal representation r_p of the relative probabilities.

In the case of a symmetric prior distribution (one in which the relative probabilities $(1-p, p)$ are exactly as likely as $(p, 1-p)$ for all p), we should have $\mu_z = 0$. Our results above then imply that p_0 should equal 0.5, and that we should observe that subjects' median bids should satisfy $|WTP| > |EV|$ for lotteries with $p < 0.5$ and $|WTP| < |EV|$ for lotteries with $p > 0.5$, in either the gain or loss domain, as Enke and Graeber (2023) find.[99]

Moreover, fixing the prior distribution of the probabilities, the size of these biases (i.e., the systematic departures from risk-neutral bidding) should depend only on σ_z^2, the degree of imprecision in the internal representation of probabilities. A larger value of σ_z^2 should increase $\hat{\sigma}_z^2$, and as a consequence should lower the value of α. It should also make $\hat{\mu}_z(z)$ closer to zero, for any value of z. Hence for both reasons, the median value $\bar{p}$ of the posterior mean estimate of p given by (B.8) should be closer to 0.5, for any true p, the larger is σ_z^2. This in turn means that for any $p \neq 0.5$, the size of the departure from risk-neutral bidding implied by (B.9) should be an increasing function of σ_z^2. This prediction is consistent both with the results of Enke and Graeber that show that subjects with higher reported cognitive uncertainty exhibit larger departures from risk-neutrality (in all four quadrants of the Tversky-Kahneman "fourfold pattern"), and with their demonstration that interventions that ought to reduce the precision of subjects' awareness of the value of p cause them to exhibit larger departures from risk-neutrality (again in all four quadrants).

We show how these predictions can be extended to the more general case in which there is cognitive uncertainty about the magnitude $|X|$ of the monetary payoff as well, and also derive the consequences of taking into account unavoidable response error, in the section that follows.

C Noisy Coding and Lottery Valuation: Derivations

Here we explain the details of the derivation of the theoretical model sketched in the main text. We begin with a complete derivation of the quantitative predictions of our baseline model, and then briefly discuss the predictions of the alternative models of noisy coding that are compared in Table 3.

C.1 Losses from Inaccurate Bidding

We begin by explaining why minimization of the expected loss (2.4) corresponds to maximization of the DM's expected financial wealth, given the financial incentives provided in our experiment. Our subjects are incentivized by conducting a BDM auction at the end of the experiment, for the lottery offered in one randomly selected trial; if the subject wins this auction (bids an amount greater than the random bid generated for their automated opponent), he receives

[99] The findings of Enke and Graeber, of course, are equally consistent with a model in which it is EV (rather than p) that is represented (or retrieved) with noise. But as discussed in section 4.2 of the main text, that alternative type of model of optimal adaptation to cognitive noise is much less successful at explaining the pattern of apparent risk attitudes in our experimental data.

the outcome of the lottery (but has his endowment reduced by the amount of the opponent's bid), while if not (because the opponent bids more), he keeps his endowment.

Let W_0 be the subject's initial wealth (inclusive of the endowment received in the experiment), N the number of trials in the experiment (each of which has a probability $1/N$ of being selected as the basis for the subject's payment), C_i the amount that the subject bids for the lottery on trial i, and (p_i, X_i) are the characteristics of the lottery offered on that trial. In the case that trial i is selected for payment, the random bid B_i of the automated opponent is an independent draw from a distribution with continuous density function $g(B)$. The mathematical expectation of the subject's wealth at the end of the experiment, conditional on their sequence of bids $\{C_i\}$, is then equal to

$$\begin{aligned} & W_0 \;+\; \frac{1}{N}\sum_i \mathrm{E}[I(B_i < C_i)\cdot(p_iX_i - B_i)] \\ = \; & W_0 \;+\; \frac{1}{N}\sum_i \mathrm{E}[I(B_i < p_iX_i)\cdot(p_iX_i - B_i)] \;-\; \frac{1}{N}\sum_i L(C_i;\, p_iX_i), \end{aligned} \tag{C.1}$$

where

$$L(C;\, V) \;\equiv\; -\int_V^C (V - B)\cdot g(B)dB. \tag{C.2}$$

Here $I(\cdot)$ is an indicator function, taking the value 1 if the statement inside the parentheses is true, and 0 otherwise; and the symbol $\mathrm{E}[\cdot]$ refers to the mathematical expectation over possible realizations of the random variables p_i, X_i, C_i, and B_i.

The first two terms in (C.1) are independent of the subject's bid. Hence a bidding rule maximizes the expectation of the subject's wealth if and only if it minimizes the final term in (C.1). Since this final term is a sum of additively separable terms for the different trials, we can consider separately the optimal bidding rule to use in a single trial. Thus an optimal bidding rule is one that chooses a bid C (or a probability distribution for such bids), given an internal representation of the decision situation $\mathbf{r}$, so as to minimize the expected loss

$$\mathrm{E}[L(C;\, pX)\,|\mathbf{r}]. \tag{C.3}$$

If we suppose that the distribution $g(B)$ is approximately uniform[100] — that $g(B) \approx \tilde{g}$ over the relevant range,[101] where $\tilde{g} > 0$ is a constant — then we can approximate (C.2) as

$$L(C;\, pX) \;\approx\; \tilde{g}\cdot\int_{pX}^C (pX - B)\, dB \;=\; \frac{\tilde{g}}{2}\, Q,$$

where $Q \equiv (C - pX)^2$. Thus as long as $g(B)$ is uniform over a sufficiently wide range, minimization of (C.3) is equivalent to minimization of the objective (2.4) assumed in the main text.[102]

[100]This is an assumption about the prior beliefs of the DM, for which the DM's bidding rule are assumed to be optimized. Note that the BDM auction in our experiment involves a uniform distribution $g(B)$.

[101]We mean, for $0 < B < \bar{B}$, where $\bar{B}$ is large enough so that $pX < \bar{B}$ with high probability, under the prior used to evaluate (C.2).

[102]The omission of the positive multiplicative factor $\tilde{g}/2$ affects only the units in which costs are measured, and not the validity of the conclusions reached in our analysis. If one wanted to interpret the estimated precision cost parameter A in terms of dollars per additional unit of encoding precision, the numerical value of $\tilde{g}$ would matter; but the fact that there should be a positive structural parameter A under the assumptions of our baseline model is independent of the value of $\tilde{g}$.

If C could be chosen with precision, given an internal representation $\mathbf{r}$, the solution to this problem would be to choose the bid specified in (2.5). That is, the optimal bid would simply be the mean of the Bayesian posterior distribution for the true expected value of the lottery, conditional on the imprecise internal representation of the problem. However, because of the presence of unavoidable response error, it is only the mean of the distribution (2.3) that can be chosen as a function of $\mathbf{r}$, and not the value of C that will be bid on any given trial. If response error were assumed to be additive, a "certainty equivalence" result would obtain: (2.5) would still be the optimal value for the intended bid, though the actual bid would equal this plus a mean-zero noise term. But because we have (more accurately, in our view) specified a multiplicative response error in (2.3), the optimal solution is more complex, as we discuss below.

C.2 Implications of Cognitive Noise for Optimal Bidding

To simplify the discussion, we first consider the case of a lottery in which there is a probability p of obtaining a positive monetary payoff X. The quantities (p, X) that specify the decision problem on a given trial have noisy internal representations (r_p, r_x), the conditional distributions of which are given by

$$r_p|p \;\sim\; N(\log(p/1-p),\, \nu_z^2), \qquad r_x|(r_p, X) \;\sim\; N(\log X,\, \nu_x^2(r_p)),$$

where the function $\nu_x^2(r_p)$ is to be optimized (but is taken as given in this section). Note that the conditional distribution of r_p is independent of the magnitude of X, and that the conditional distribution of r_x depends on the value of p only through its internal representation r_p. We can view r_p as being determined first, in a way that depends only on the value of p; the internal representation r_x is then determined by X, but in a way that can depend on the already encoded value r_p.

The DM's optimal bid as a function of the internal representation (r_p, r_x) depends on the prior distribution from which the true values (p, X) are expected to have been drawn. We suppose that p and X are independent random variables, with a prior distribution for X given by

$$\log X \;\sim\; N(\mu_x,\, \sigma_x^2).$$

(The conclusions in this section are independent of what we assume about the prior distribution for p, other than that the two variables are distributed independently of one another.) Under the assumption of a log-normal prior for X, the posterior for X is also log-normal. It follows that

$$\mathrm{E}[X\,|r_p, r_x] \;=\; \exp[(1-\gamma_x(r_p))\mu_x \;+\; \gamma_x(r_p) r_x \;+\; \frac{1}{2}(1-\gamma_x(r_p))\sigma_x^2],$$

where

$$\gamma_x(r_p) \;\equiv\; \frac{\sigma_x^2}{\sigma_x^2 + \nu_x^2(r_p)}$$

is a quantity satisfying $0 < \gamma_x(r_p) < 1$ that can be different for each r_p. It similarly follows that

$$\mathrm{E}[X^2\,|r_p, r_x] \;=\; \exp[2(1-\gamma_x(r_p))\mu_x \;+\; 2\gamma_x(r_p) r_x \;+\; 2(1-\gamma_x(r_p))\sigma_x^2].$$

As explained in the main text, we assume that the DM's bid C on a given trial is drawn from a distribution of possible bids

$$\log C \;\sim\; N(f(r_p, r_x),\, \nu_c^2),$$

where the function $f(r_p, r_x)$ is to be optimized. Note that we can alternatively write

$$\log C \;=\; f(r_p, r_x) \,+\, \epsilon_c, \tag{C.4}$$

where

$$\epsilon_c \;\sim\; N(0,\, \nu_c^2)$$

is distributed independently of r_p and r_x.

We now consider the optimal choice of the function f. For each possible internal representation (r_p, r_x), we have a separate optimization problem: choose $f(r_p, r_x)$ to minimize

$$\begin{aligned}
\mathrm{E}[Q\,|r_p, r_x] \;&=\; \mathrm{E}[C^2\,|r_p, r_x] \,-\, 2\mathrm{E}[CpX\,|r_p, r_x] \,+\, \mathrm{E}[p^2X^2\,|r_p, r_x] \\
&=\; \mathrm{E}[\exp(2\epsilon_c)] \cdot \exp(2f(r_p, r_x)) \\
&\quad -\, 2\mathrm{E}[\exp(\epsilon_c)] \cdot \exp(f(r_p, r_x)) \cdot \mathrm{E}[p\,|r_p] \cdot \mathrm{E}[X\,|r_p, r_x] \\
&\quad +\, \mathrm{E}[p^2\,|r_p] \cdot \mathrm{E}[X^2\,|r_p, r_x],
\end{aligned}$$

where we have used (C.4) to substitute for C as a function of r_p, r_x, and ϵ_c.

This is a quadratic function of $\exp(f(r_p, r_x))$. Moreover, since

$$\mathrm{E}[\exp(2\epsilon_c)] \;=\; \exp(2\nu_c^2) \;>\; 0,$$

it is a strictly convex function, with a unique minimum when

$$\mathrm{E}[\exp(2\epsilon_c)]\exp(f(r_p, r_x)) \;=\; \mathrm{E}[\exp(\epsilon_c)] \cdot \mathrm{E}[p\,|r_p] \cdot \mathrm{E}[X\,|r_p, r_x].$$

Using the fact that both X and ϵ_c are log-normally distributed (conditional on r_p, r_x), we can express the optimal choice of f as

$$f(r_p, r_x) \;=\; \log \mathrm{E}[p\,|r_p] \,+\, (1 - \gamma_x(r_p))[\mu_x + \frac{1}{2}\sigma_x^2] \,+\, \gamma_x(r_p)r_x \,-\, \frac{3}{2}\nu_c^2.$$

When f is chosen in this way, the minimized value of the quadratic function is

$$\begin{aligned}
\mathrm{E}[Q\,|r_p, r_x] \;&=\; \exp(2(1 - \gamma_x(r_p))[\mu_x + \frac{1}{2}\sigma_x^2] \,+\, 2\gamma_x(r_p)r_x) \cdot \\
&\qquad \big\{\exp((1 - \gamma_x(r_p))\sigma_x^2)\mathrm{E}[p^2\,|r_p] \,-\, \exp(-\nu_c^2)\mathrm{E}[p\,|r_p]^2\big\}.
\end{aligned} \tag{C.5}$$

Substitution of this solution into (C.4) implies that the equation

$$\begin{aligned}
\log C \,-\, \log(pX) \;&=\; \big(\log \mathrm{E}[p\,|r_p] - \log p\big) \,+\, (1 - \gamma_x(r_p))[\mu_x + \frac{1}{2}\sigma_x^2 - \log X] \\
&\qquad + \gamma_x(r_p)[r_x - \log X] \,-\, \frac{3}{2}\nu_c^2 \,+\, \epsilon_c
\end{aligned}$$

gives the predicted value of $\log(WTP/EV)$ in the case of any given lottery (p, X), any given internal representation (r_p, r_x), and any given realization of the response noise ϵ_c. Integrating over the conditional distributions of the random variables (r_p, r_x, ϵ_c) in the case of a given lottery (p, X), we obtain the prediction that

$$\mathrm{E}[\log(C/pX)\,|p, X] \;=\; \alpha_p \;+\; \beta_p \log X, \tag{C.6}$$

where the coefficients

$$\alpha_p \;\equiv\; \mathrm{E}\big[\log \mathrm{E}[p\,|r_p] - \log p\;|p\big] \;+\; (1-\gamma_p)[\mu_x + \frac{1}{2}\sigma_x^2] \;-\; \frac{3}{2}\nu_c^2, \tag{C.7}$$

$$\beta_p \;\equiv\; -(1-\gamma_p),$$

$$\gamma_p \;\equiv\; \mathrm{E}[\gamma_x(r_p)\,|p]$$

all depend on the value of p but are independent of X. (Note that, among other things, this solution implies equation (3.9) in the main text.)

Since $0 < \gamma_x(r_p) < 1$ for each possible value of r_p, it follows that $0 < \gamma_p < 1$ for each value of p, and hence that $-1 < \beta_p < 0$ for each p. We thus conclude that for any lottery (p, X), the predicted distribution of values for WTP (i.e., the distribution of the random variable C in (C.6)) is such that the mean value of $\log(WTP/EV)$ should be an affine function of $\log X$, with a slope and intercept that can vary with p. Furthermore, for each value of p, the slope must satisfy $-1 < \beta_p < 0$. These predictions are tested in the way discussed in section 1.3 of the main text.

In the case that X is negative (the lottery offers a random loss rather than a random gain), we suppose that p and the magnitude $|X|$ are encoded with noise in the same way (and with the same parameters) as is specified above in the case that X is positive. The optimal bid in this case will obviously be negative; we assume that in the case of a negative bid C, the absolute value $|C|$ will again be given by the right-hand side of (C.4), just as in the case of a positive bid. The optimal function $f(r_p, r_x)$ will then be exactly the same as in the derivation above. We conclude that the distribution of values for C/pX will be exactly the same function of p and $|X|$ as in the case where X is positive. In particular, (C.6) will again hold, except with $\log X$ replaced by $\log |X|$ on the right-hand side; the coefficients α_p, β_p will be the same functions of p as in the case of random gains. This prediction is also tested in the way discussed in the main text.

C.3 Endogenous Encoding Precision

We turn now to the way in which the coefficients α_p, β_p are predicted to vary with p. This depends on what we assume about the noisy encoding of p, and about the prior over values of p for which the decision rule is optimized; but it also depends on what we assume about how $\nu_x^2(r_p)$ varies with r_p. We suppose that the latter function is endogenously determined, so as to maximize the accuracy of bidding subject to a cost of encoding precision, as discussed in the main text.

Note that our model of noisy coding implies that conditional on the value of r_p, the distribution of r_x is

$$r_x|r_p \;\sim\; N(\mu_x,\, \sigma_x^2 + \nu_x^2(r_p)),$$

from which it follows that

$$2\gamma_x(r_p)r_x|r_p \;\sim\; N(2\gamma_x(r_p)\mu_x,\, 4\gamma_x(r_p)\sigma_x^2).$$

Thus exponentiation of this variable results in a log-normal random variable, with mean

$$\mathrm{E}[\exp(2\gamma_x(r_p)r_x)\,|r_p] \;=\; \exp(2\gamma_x(r_p)\mu_x \,+\, 2\gamma_x(r_p)\sigma_x^2).$$

Using this result, we can then integrate (C.5) over the possible realizations of r_x to obtain

$$\mathrm{E}[Q\,|r_p] \;=\; \frac{\tilde{g}}{2}\cdot\exp(2(\mu_x+\frac{1}{2}\sigma_x^2))\cdot\left\{\exp(\sigma_x^2)\mathrm{E}[p^2\,|r_p] \,-\, \exp(\gamma_x(r_p)\sigma_x^2-\nu_c^2)\mathrm{E}[p\,|r_p]^2\right\}.$$

Thus we can write

$$\mathrm{E}[Q\,|r_p] \;=\; Z(r_p) \,-\, \Gamma\varphi(r_p)\cdot\exp(\gamma_x(r_p)\sigma_x^2),$$

where

$$\Gamma \;\equiv\; \frac{\tilde{g}}{2}\exp(-\nu_c^2) \;>\; 0, \qquad \varphi(r_p) \;\equiv\; \mathrm{E}[p\,|r_p]^2 \;>\; 0,$$

and $Z(r_p)$ are all positive factors with values that are independent of the choice of $\nu_x^2(r_p)$. We thus observe that for any r_p, the value of $\mathrm{E}[\tilde{L}\,|r_p]$ is a monotonically decreasing function of $\gamma_x(r_p)$, and hence a monotonically increasing function of $\nu_x^2(r_p)$.

If the cost of greater precision in the encoding of X, in units of money (i.e., the same units as those in which $L(C;pX)$ is measured), is given by $(\tilde{g}/2)\kappa(\nu_x^2)$, where

$$\kappa(\nu_x^2) \;=\; \frac{\tilde{A}}{\nu_x^2} \;=\; \frac{\tilde{A}}{\sigma_x^2}(\frac{\gamma_x}{1-\gamma_x}),$$

then minimization of total costs (counting the cost of precision) requires that for each r_p, the value of $\gamma_x(r_p)$ be the solution to the problem

$$\min_{\gamma_x} F(\gamma_x;\, r_p) \;\equiv\; \frac{\tilde{A}}{\sigma_x^2}(\frac{\gamma_x}{1-\gamma_x}) \;-\; \Gamma\varphi(r_p)\cdot\exp(\gamma_x\sigma_x^2). \tag{C.8}$$

We further observe that

$$\frac{\partial F}{\partial\gamma_x} \;=\; \frac{\tilde{A}}{\sigma_x^2}\frac{1}{(1-\gamma_x)^2} \;-\; \Gamma\varphi(r_p)\sigma_x^2\cdot\exp(\gamma_x\sigma_x^2),$$

an expression that has a positive sign if and only if

$$\frac{A}{(1-\gamma_x)^2} \;>\; \varphi(r_p)\exp(\gamma_x\sigma_x^2), \tag{C.9}$$

where we now use

$$A \;\equiv\; \frac{\tilde{A}}{\Gamma\sigma_x^4} \;>\; 0$$

as an alternative parameterization of the size of the cost of precision. Taking the logarithm of both sides of the inequality (C.9), we see that

$$\frac{\partial F}{\partial \gamma_x} \;>\; 0 \;\Leftrightarrow\; G(\gamma_x;\, r_p) \;>\; 0,$$

where we define

$$G(\gamma_x;\, r_p) \;\equiv\; \log A \;-\; \log \varphi(r_p) \;-\; 2\log(1-\gamma_x) \;-\; \gamma_x \sigma_x^2. \tag{C.10}$$

We see from this that $F(\gamma_x;\, r_p)$ is a decreasing function of γ_x at $\gamma_x = 0$ if and only if

$$A \;<\; \varphi(r_p), \tag{C.11}$$

so that $G(0;\, r_p) \;<\; 0$. We also note that $F(\gamma_x;\, r_p)$ is an increasing function of γ_x as $\gamma \to 1$ (indeed, increasing without bound). Hence (C.11) is a sufficient condition for the existence of an interior solution to the problem (C.8) at some $0 < \gamma_x < 1$. Moreover, the function defined in (C.10) is a strictly convex function of γ_x; hence its graph can cross the line $G = 0$ for at most two values of γ_x, and then only if $G > 0$ at both extremes.

Thus if (C.11) holds, there must be exactly one solution to the first-order condition

$$G(\gamma_x;\, r_p) \;=\; 0, \tag{C.12}$$

an equivalent way of writing condition (3.8) stated in the main text. (Condition (3.8) in the main text is just the requirement that (C.9) hold as an equality.) In addition, we must have $G < 0$ for all smaller values of γ_x, while $G > 0$ for all greater values of γ_x. From this it follows that the solution to the FOC must be the global minimum of the function F, and hence the solution to problem (C.8).

We also observe that the value of r_p affects this solution only through its effect on the value of $\varphi(r_p)$; thus we can solve for the optimal γ_x as a function of the value of $\varphi(r_p)$. When $\varphi(r_p)$ satisfies (C.11), so that we have an interior solution to the FOC, we can compute the derivative of γ_x with respect to changes in the value of $\varphi(r_p)$ through total differentiation of the FOC. It follows from (C.10) that

$$\frac{\partial G}{\partial \varphi} \;=\; -\frac{1}{\varphi} \;<\; 0, \qquad \frac{\partial G}{\partial \gamma_x} \;=\; \frac{2}{1-\gamma_x} \;-\; \sigma_x^2 \;>\; 0,$$

if $\sigma_x^2 \leq 2$ as assumed in the main text. Then total differentiation of the FOC (C.12) implies that

$$\frac{d\gamma_x}{d\varphi(r_p)} \;=\; -\frac{\partial G/\partial \varphi}{\partial G/\partial \gamma_x} \;>\; 0.$$

It follows that the optimal solution for γ_x will be a monotonically increasing function of $\varphi(r_p)$, with $\gamma_x \to 0$ as $\varphi \to A$ and $\gamma_x \to 1$ as $\varphi \to \infty$. Or equivalently, the optimal solution for ν_x^2 will be a monotonically decreasing function of $\varphi(r_p)$, with $\nu_x^2 \to \infty$ as $\varphi \to A$ and $\nu_x^2 \to 0$ as $\varphi \to \infty$.

Let us now consider the alternative case in which $\varphi(r_p) \leq A$. In this case $G \geq 0$ when $\gamma_x = 0$, and since $\partial G/\partial \gamma_x > 0$ (again assuming that $\sigma_x^2 \leq 2$), it follows that $G > 0$ for all $\gamma_x > 0$. This implies that $\partial F/\partial \gamma_x > 0$ for all $\gamma_x > 0$, so that the solution to the problem (C.8) must be $\gamma_x = 0$ in all such cases. Thus we obtain a unique optimal solution for γ_x (and hence for ν_x^2) for any value of $\varphi(r_p)$. The optimal γ_x is a non-decreasing function of $\varphi(r_p)$: constant (and equal to zero) for all $0 \leq \varphi(r_p) \leq A$, and increasing for all $\varphi(r_p) > A$.

C.4 Alternative Models of Noisy Coding

In section 4.2 of the main text, we consider a variety of alternatives to the baseline cognitive noise model analyzed above. Here we briefly discuss how the quantitative predictions of each of these models are derived.

Exogenous precision. This model assumes as a constraint that $\nu_x(r_p) = \nu_x$ for all r_p, where ν_x is now a constant to be estimated. In this case, the expressions derived above continue to apply, but γ_x is now a quantity (between 0 and 1) independent of r_p. As a result, γ_p takes the same value (equal to γ_x) for all p, and its value can be calculated from the values of parameters σ_x and ν_x, independently of assumptions about the encoding and decoding of the probabilities. An important implication is that $\beta_p = -(1-\gamma_p)$ will be the same negative quantity for all values of p; thus stake-size effects are predicted to be the same for all values of p (contrary to what we, and others, find empirically).

Noiseless retrieval of monetary payoffs. This model assumes that $\nu_x = 0$; it is thus a special case of the exogenous noise model, in which ν_x is no longer a free parameter. In this special case, we have the stronger result that $\gamma_x = 1$ (again regardless of the value of r_p), and hence that $\beta_p = 0$ for all p. Thus this model implies that there should be no stake-size effects.

Noiseless retrieval of probabilities. This model assumes that $\nu_z = 0$, so that r_p can be identified with the value of p itself. In this case, condition (3.8) reduces to the equation

$$\frac{A}{(1-\gamma_p)^2} \;=\; p^2 \cdot \exp(\gamma_p \sigma_x^2).$$

This is an equation that implicitly defines the value of γ_p (and hence the value of β_p) for any value of p; it is no longer necessary to solve for $\gamma_x(r_p)$ for each of the possible r_p associated with a given probability p, and then numerically integrate over the distribution of such solutions in order to obtain a prediction for γ_p. Condition (C.7) yields an expression for α_p that can be solved in closed form, and that is the same for all p. And finally, we obtain a closed-form solution for the variance of the distribution of $\log(WTP)$ as well, namely

$$\mathrm{var}(\log(WTP)\,|p, X) \;=\; \gamma_p^2 \nu^2 \,+\, \nu_c^2 \;=\; \gamma_p(1-\gamma_p)\sigma_x^2 \,+\, \nu_c^2.$$

Thus obtaining accurate numerical predictions for the distribution of bids is much simpler in this case than in the case of the baseline model.

Among the implications that follow in this special case: for all values of p, one obtains the prediction that

$$\mathrm{E}[\log(WTP/EV)\,|p, X] \;=\; -\frac{3}{2}\nu_c^2$$

when X is equal to its prior mean. This means that when X is equal to its prior mean, varying p cannot change the mean $\log(WTP)$; and since (2.3) implies that $\log(WTP)$ is necessarily a Gaussian distribution with variance ν_c^2, it follows that the entire distribution of values for $\log(WTP/EV)$ must be independent of p. Thus the sign of the relative risk premium cannot change as we vary p (as asserted by the fourfold pattern of risk attitudes of Kahneman and Tversky), and indeed not even its magnitude can change. It is for this reason that this case is quite inconsistent with our data.

No response noise. This model assumes that $\nu_c = 0$, so that the DM can directly (and optimally) choose their bid C as a function of the noisy internal representation $\mathbf{r}$. In this case, the optimal bidding rule is simply $C = \mathrm{E}[pX\,|\mathbf{r}]$, as assumed in the discussion in the introduction. The formulas stated above continue to hold, but with the simplification that terms involving ν_c can be omitted.

Noisy encoding and retrieval of EV. In this model, we assume (by analogy with (refrxdist)) that the internal representation of $|EV|$ is drawn from a distribution

$$r_{ev} \;\sim\; N(\log|EV|,\, \nu_{ev}^2),$$

where the variance ν_{ev}^2 is independent of the lottery's EV. The DM's bid is then assumed to have the sign of the EV (which is recognized with perfect accuracy), and a magnitude that is drawn from a distribution

$$\log|C| \;\sim\; N(f(r_{ev}),\, \nu_c^2), \tag{C.13}$$

by analogy with (2.3). The bidding function $f(r_{ev})$ is again determined so as to minimize the objective (2.4). Computing the value of this objective requires that we specify a prior regarding the true values of the lotteries with which the DM may be presented. We suppose that the DM bidding rule is optimized for a log-normal prior,

$$\log|EV| \;\sim\; N(\mu_{ev},\, \sigma_{ev}^2),$$

where (just as in the case of our other cognitive noise models) the parameters (μ_{ev}, σ_{ev}) of the prior are the ones that maximize the likelihood of the values of EV actually used in the experiment.

For the same reason as in our derivation for the baseline model, the optimal bidding function will be of the form

$$f(r_{ev}) \;=\; \log \mathrm{E}[|EV|\,|r_{ev}] \;-\; \frac{3}{2}\nu_c^2. \tag{C.14}$$

And as above, we can use the algebra of log-normal distributions to show that under these assumptions, the posterior mean will be given by

$$\mathrm{E}\Big[|EV|\,|r_{ev}\Big] \;=\; \exp\Big((1-\gamma_{ev})\bar{\mu}_{ev} \;+\; \gamma_{ev}\cdot r_{ev}\Big), \tag{C.15}$$

where we define

$$\gamma_{ev} \;\equiv\; \frac{\sigma_{ev}^2}{\sigma_{ev}^2 + \nu_{ev}^2} \;<\; 1, \qquad \bar{\mu}_{ev} \;\equiv\; \mu_{ev} \;+\; \frac{1}{2}\sigma_{ev}^2.$$

Equations (C.13)–(C.15) then completely specify the predicted log-normal distribution of bids implied by any internal representation r_{ev}.

From this, we obtain the prediction that $m(p,X)$ should be a log-linear function of the form (1.2), with

$$\alpha_p \;=\; (1-\gamma_{ev})[\bar{\mu}_{ev} - \log p] \;-\; \frac{3}{2}\nu_c^2, \qquad \beta_p \;=\; -(1-\gamma_{ev}),$$

and that

$$v(p,X) \;=\; \gamma_{ev}^2\nu_{ev}^2 \;+\; \nu_c^2$$

for all (p, X). Thus this model, like the baseline model, predicts that $m(p, X)$ should be an affine function of $\log |X|$, with a slope between 0 and -1, that is independent of the sign of X. It also predicts that α_p should be monotonically decreasing as a function of p, rising sharply for the smallest values of p, in qualitative accordance with our estimated coefficients for the atheoretical bounded symmetric affine model. However, like the model with exogenous noise (or the model with noiseless retrieval of monetary payoffs), it predicts that β_p should be the same for all p, rather than becoming more negative for small values of p, as in our data. And it predicts a sharper rate of increase in α_p for small values of p than does the baseline model: this model predicts that α_p should equal a constant minus $\log p$, while the baseline predicts that it should equal a constant plus $\mathrm{E}[\log \mathrm{E}[p|r_p]\,|p]$ minus $\log p$. Finally, this model predicts that the trial-to-trial variability of bids (in percentage terms) should be independent of both p and $|X|$; this means that (unlike the baseline model) it fails to capture the way in which bids are more variable for low values of p.

D Maximum Likelihood Parameter Estimation

D.1 Likelihood of the Data under Alternative Models

Let y_i be the observed value on any trial i of the variable $\log(WTP/EV)$. The log-likelihood of the data $\{p_i, X_i, y_i\}$ can be expressed in the form

$$\mathrm{LL} \;=\; \sum_i \,[L_1(p_i, X_i) \;+\; L_2(y_i\,|p_i, X_i)], \tag{D.1}$$

where the sum is over the trials in the data set, indexed by i. For each trial, the contribution $L_1(p_i, X_i)$ is the log of the likelihood of the subject's being presented with lottery (p_i, X_i) according to the prior; and $L_2(y_i\,|p_i, X_i)$ is the log of the conditional likelihood of the (scaled) response y_i, given lottery (p_i, X_i), under a given parametric model of bidding behavior. In our atheoretical models, the parts L_1 and L_2 are each functions of different sets of parameters: the parameters of the priors matter only for L_1, while the behavioral parameters matter only for L_2. But in our optimal bidding model, instead, the conditional likelihoods L_2 also involve the parameters of the prior, in the way explained in Appendix section C.

We can write (D.1) in the form

$$\mathrm{LL} \;=\; \sum_j N_j L_j, \tag{D.2}$$

where the sum is over the different lotteries (indexed by j) used in the experiment, N_j is the number of trials involving lottery j, and L_j is the average contribution to the log likelihood from the trials involving that lottery. Each term L_j depends only on the data for trials $i \in I_j$, the set of trials on which $(p_i, X_i) = (p_j, X_j)$. Thus L_j depends only on p_j, X_j, and the bids $\{WTP_i\}$ for trials $i \in I_j$. We can also further decompose each of the terms LL_j in the same way as in (D.1), writing

$$L_j \;=\; L_1(p_j, X_j) \;+\; L_{2,j}, \tag{D.3}$$

where

$$L_{2,j} \;=\; \frac{1}{N_j}\sum_{i\in I_j} L_2(y_i\,|p_j, X_j).$$

The L_1 terms are the same for all of the models that we consider in this paper. Our specifications (2.8) and (2.9) for the prior imply that

$$L_1(p_j, X_j) \;=\; -\frac{1}{2}\left(\frac{\log|X_j| - \mu_x}{\sigma_x}\right)^2 \;-\; \log(\sqrt{2\pi}\sigma_x) \;-\; \log(2\sqrt{3}\sigma_z), \tag{D.4}$$

for any p_j such that

$$\mu_z - \sqrt{3}\sigma_z \;\le \log\frac{p_j}{1-p_j} \;\le\; \mu_z + \sqrt{3}\sigma_z. \tag{D.5}$$

(Here we have omitted certain additive terms in (D.4) that are independent of the assumed parameter values; these terms have no effect on our judgments about the relative value of LL under different parameter values, and hence no effect on our maximum-likelihood parameter estimates or our model-comparison statistics.)

If p_j instead falls outside the interval (D.5), i.e., outside the support of the prior (2.9), given the assumed parameter values, then the prior probability of such an observation is zero, and $L_1(p_j, X_j) = -\infty$. Hence in our search for maximum-likelihood parameter values, we can impose as a constraint that the parameters of the prior must satisfy

$$\mu_z - \sqrt{3}\sigma_z \;\le \min_j \log\frac{p_j}{1-p_j}, \qquad \mu_z + \sqrt{3}\sigma_z \;\ge \max_j \log\frac{p_j}{1-p_j},$$

where the minimum and maximum are over the set of probabilities used in the experiment.[103] Subject to these constraints, we find values of the parameters that maximize the function LL, using expression (D.4) for the L_1 terms.

In each of the atheoretical characterizations of the data considered in Table 1, we assume a distribution of bids for the lottery (p_j, X_j) of the form

$$y_i \;\sim\; N(m_j,\, v_j) \tag{D.6}$$

on each trial $i \in I_j$; the models differ only in the restrictions that they place on the possible values of the parameters $\{m_j,\, v_j\}$. In the case of any model of this kind, the average contribution of each trial involving lottery j to the conditional log-likelihood of the data is then given by

$$L_{2j} \;=\; -\frac{1}{2v_j}\left[\hat{v}_j \,+\, (\hat{m}_j - m_j)^2\right] \;-\; \frac{1}{2}\log(2\pi v_j), \tag{D.7}$$

where we define the sample mean and variance of the data as

$$\hat{m}_j \;\equiv\; \frac{1}{N_j}\sum_{i\in I_j} y_i, \qquad \hat{v}_j \;\equiv\; \frac{1}{N_j}\sum_{i\in I_j}(y_i - \hat{m}_j)^2.$$

[103] As shown in Appendix section E.1, these minimum and maximum probabilities are 0.05 and 0.95 respectively.

Note that in (D.7), the quantities m_j, v_j are parameters of the model (the values of which are estimated to fit the data), while the quantities $\hat{m}_j, \hat{v}_j$ are data moments. Given the data, the MLE estimates for the parameters (in the absence of any further restrictions) will depend only on these moments of the data, and are equal to[104]

$$m_j \;=\; \hat{m}_j, \qquad v_j \;=\; \hat{v}_j.$$

Thus in the case of any model parameters $\{m_j,\, v_j\}$, the value of the log-likelihood LL can be computed from the data moments $\{\hat{m}_j,\, \hat{v}_j\}$, using equations (D.2) – (D.4) and (D.7).

In the results reported in Table 1, the parameters of each of the various atheoretical statistical models are fit to the data moments of a fictitious "average subject." For each lottery j, we define $\hat{m}_j^{avg}$ as the median value of $\hat{m}_j$ across the various subjects who bid on lottery j, and $\hat{v}_j^{avg}$ as the median value of $\hat{v}_j$ across these same subjects. (These are the data moments plotted in Figures 2 and 3.) In order to compute the log likelihood for any model parameters $\{m_j,\, v_j\}$ using equations (D.2) – (D.4) and (D.7), using $\{\hat{m}_j^{avg},\, \hat{v}_j^{avg}\}$ for the data moments in (D.7). For the quantity N_j in (D.2), we use N_j^{avg}, the effective number of observations of bids on lottery j by the average subject. This is defined as

$$N_j^{avg} \;\equiv\; \frac{1}{H_j} \sum_h N_j^h,$$

where H_j is the number of subjects bidding on lottery j.[105]

The MLE estimates of the parameters of the cognitive noise models are chosen in a similar way, to maximize the log likelihood of the average-subject data. The exact solution to the optimal bidding model does not imply that a DM's bids on a given lottery should be drawn from a log-normal distribution, as specified in (D.6); while (C.4) implies a log-normal distribution of bids conditional on the internal representation $\mathbf{r}$, when we condition on the true lottery characteristics (as in our computation of the data moments) rather than on the unobserved internal representation, the predicted distribution should instead be a mixture of log-normal distributions. For purposes of model fitting, however, we use a Gaussian approximation to the model predictions, according to which y_i should have a log-normal distribution as specified in (D.6), the parameters of which are given by the mean and variance of $\log y_i$ predicted by the optimizing model. Using this approximation, we can compute an approximate likelihood of the data under any assumed model parameters, simply on the basis of data for the first and second moments $\{\hat{m}_j^{avg},\, \hat{v}_j^{avg}\}$.[106]

The MLE estimates of the parameters of our various theoretical models are also obtained by maximizing an approximate likelihood function calculated in this way. The reported values of LL and BIC are then based on the maximized value of the approximate likelihood function. Finally, the value of LL/N reported in Table 4 for the cognitive noise model with

[104] This explains our notation for the data moments: $\hat{m}_j$ is the MLE estimate of the parameter m_j, and $\hat{v}_j$ is the MLE estimate of the parameter v_j.

[105] Note that this is not the same for all lotteries j. The value of H_j varies between 5 (in the case of lotteries with $p = 0.3$ or 0.7) and 22 (in the case of lotteries with $p = 0.1$ or 0.9); see Table 6 below.

[106] Use of this approximation is desirable, not simply as a way of simplifying our numerical solution for the likelihood, but because we have only defined the first and second moments of the "average subject data" — we don't have a complete sample of bids by the fictitious "average subject."

Variants of the Baseline Model			
model	A	ν_z	ν_c
baseline model	0.002	1.60	0.24
no payoff noise	0	1.61	0.25
no probability noise	0.019	0	0.50
no response noise	0.004	1.75	0
Model with Exogenous Precision			
	ν_x	ν_z	ν_c
	0.16	1.61	0.24
Model with Noisy Retrieval of EV			
	ν_{ev}		ν_c
	0.95		0.24

Table 5: Parameter estimates for the alternative cognitive noise models, for which model comparison statistics are given in Table 3 of the main text. For the interpretation of the reported parameters, see the explanation of these models in Appendix section C.4.

a single average subject actually divides LL by $N^{avg} = N/H$, the average number of bids per subject, where N is the total number of trials on which non-zero bids are submitted and H is the number of subjects in the group from which we select the median moments.

The same method is used in Table 4 to compute MLE parameter estimates (and values for LL and BIC) based on the data for other "average subjects." For example, in the case of the 640-trial average subject, the lotteries j for which the moments are computed are only the 80 lotteries used for subjects in group 5 (the 640-trial subjects), and the sums are only over the subjects h that belong to group 5.[107] In (D.2), N_j is now understood to mean $\sum_h N_j^h$, where the sum is only over the subjects in group 5. Finally, in calculating N_j^{avg}, we use the number of subjects in the 640-trial group for the value of H_j;[108] and in computing LL$/N$, we use a value N^{avg} that divides the total number of trials by the 640-trial subjects by the number of such subjects.[109]

In the case of the 400-trial average subject, we similarly compute moments only for the 100 lotteries that are evaluated by at least some of the subjects in groups 1-4 (the 400-trial subjects), and for each lottery j of this kind, we sum only over the subjects h in the groups that evaluate lottery j. For each lottery j in this set of 100 lotteries, H_j is the number of 400-trial subjects who evaluate lottery j (which varies across lotteries). And in computing LL$/N$, we use a value N^{avg} that divides the total number of non-zero bids by the 400-trial subjects by the number of such subjects.[110]

[107] For the different groups of subjects, and the lotteries evaluated by each group, see Appendix section E.1 below.

[108] Thus $H_j = 12$ in the case of the 640-trial average subject, for each of the lotteries on which group 5 bid.

[109] Note that the number of bids N^{avg} by the 640-trial average subject is not 640, because of the trials on which subjects in group 5 decline to bid, as discussed further in Appendix section E.2. For the 640-trial average subject, N^{avg} is actually only equal to 629.33. This is why in Table 4, the number given for LL$/N$ is not equal to the number given for LL divided by 640.

[110] Because the fraction of zero bids is smaller in the case of the 400-trial subjects, as discussed below, this results in $N^{avg} = 398.85$, a number only slightly less than 400.

group	members	number of trials	values of p
1	1-6	400	0.1, 0.4, 0.6, 0.8, 0.9
2	13-15	400	0.1, 0.3, 0.5, 0.7, 0.9
3	16	400	0.05, 0.1, 0.5, 0.9, 0.95
4	17-19	400	0.05, 0.3, 0.5, 0.7, 0.95
5	7-12, 20-28	640	0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95

Table 6: Number of trials and values of p used for different groups of subjects.

D.2 Parameter Estimates for Alternative Cognitive Noise Models

The maximum-likelihood estimated parameter values for our baseline model are reported in Table 4 in the main text, for each of a variety of choices regarding the "average subjects" whose data moments are to be fit. In Table 5, we report the corresponding parameter estimates for the alternative models of cognitive noise discussed in section 4.2. The table also repeats the parameter estimates for the baseline model (with all three types of cognitive noise, and endogenous precision of encoding of monetary payoffs), for purposes of comparison with the other models. In all cases, the model predictions are fit to the moments of the average subject whose bid distributions are reported in Figures 2-3.

E Experimental Data: Additional Details

Here we offer additional details about the data that we fit to the alternative models discussed in the main text.

E.1 Probabilities Used in the Lotteries

As explained in the main text, each subject was asked to evaluate a set of lotteries (p, X), where both p and X are drawn from a finite set of possibilities. Each of the finite set of values for p (for that subject) was paired with each of the finite set of values for X, and each of the pairs (p, X) that occurred for a given subject were presented equally often (8 times over the course of the session). The different lotteries (p, X) were presented in a random order.

However, the finite set of values p that were used was different for different groups of subjects, as indicated in Table 6. The subjects are classified in the table as members of one or another of five groups, according to the set of lotteries presented to them. (One group, group 3, consists of only a single subject, subject 16.) In the main text, we classify subjects into two larger groups, the 400-trial subjects (the union of groups 1-4 in Table 6) and the 640-trial subjects (group 5). Note that while the 640-trial subjects all faced the same set of lotteries, the 400-trial subjects did not; each of these evaluated a set of lotteries using only five values of p, but the values of p used were different across the four groups of 400-trial subjects. We do not, however, estimate separate model parameters for the individual groups of 400-trial subjects, given that (at least in the case of groups 2, 3, and 4) there are only a few subjects in each group.

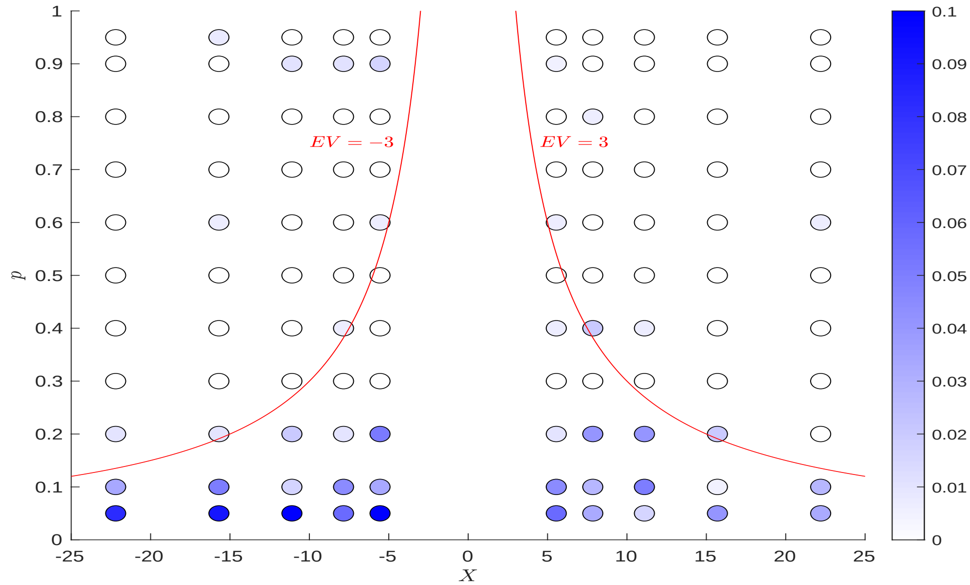

Figure 8: The fraction of zero bids for each of the lotteries (X, p) that are presented to subjects. (Color code is explained by the scale at the right.)

E.2 Zero Bids

When fitting our theoretical model to the experimental data, we exclude the bids which are equal to zero (the leftmost possible position of the slider), since, as explained in the main text, we regard this as declining to bid on that lottery. Here we provide additional information about the occurrence of these zero bids.

Zero bids were more common among the subjects in the 640-trial group (who, as noted above, also displayed more signs of inattentiveness in other respects). The 12 non-excluded 640-trial subjects submitted zero bids on a total of 160 trials, or about 1.7 percent of all trials. Zero bids were instead relatively rare for the 400-trial subjects, who submitted only 15 such bids (less than 0.3 percent of their trials).

Zero bids also occurred much more frequently for some lotteries than for others, as shown by the "heat map" in Figure 8. Zero bids are most likely to occur when p or X (or both) are small. As the figure illustrates, most of the zero bids were submitted for lotteries with an EV of less than 3 dollars (in absolute value), meaning that the optimal bid would have been in the left-most 10 percent of the range of the slider. Many are in cases where the EV is not much more than a dollar (in absolute value). Zero bids were also somewhat more common in the case of lotteries involving losses: 60 percent of the zero bids occur in these cases, even though an equal number of lotteries involving losses and gains were presented to the subjects.[111] Zero bids were especially common in the case of lotteries involving losses and

[111] This represents a departure from the symmetry of behavior in the gain and loss domains to which our data on non-zero bids by the non-excluded subjects largely conform. For example, we have shown in Table 1 that if the BIC is used as a basis for model comparison, the symmetric model is preferred to the unrestricted model, and the symmetric affine model is similarly preferred to the general affine model. But another suggestion that lotteries involving losses are more difficult to value, at least for some subjects, is the fact that three experimental subjects (subjects 9, 11 and 19) all seemed to have considerable difficulty understanding how to bid in the case of lotteries involving losses (their bids failed to increase monotonically

only a small probability ($p = 0.05$) of a non-zero loss; in this case, zero bids were submitted on 6.8 percent of all trials.

We assume that the decision whether to bother to submit a (non-zero) bid is based on a cursory inspection of the terms of the lottery (p, X). This can be modeled as a decision rule conditioned on some noisy internal representation of the information (p, X), though the information used for this first-stage decision need not be the same internal representations (r_p, r_x) that are used to choose a non-zero bid in the second stage (when it is reached). After all, we suppose that declining to submit a bid allows a saving of cognitive effort of some kind; this might mean not having to retrieve the noisy representations (r_p, r_x) that are instead needed if the DM chooses to submit a bid.[112]

Given the first-stage noisy internal representation and the first-stage decision rule, a DM has some probability $s(p, X)$ of choosing to submit a non-zero bid on a trial when the lottery is (p, X).[113] The DM's prior in the second stage (when it is reached) should then depend on this selection effect. If $\pi(p, X)$ represents the distribution from which the experimenter draws values of (p, X), then the DM's second-stage prior should be given by

$$\tilde{\pi}(p, X) \;=\; \frac{\pi(p, X)s(p, X)}{\mathrm{E}_{\pi}[s]}.$$

However, we simply take the second-stage prior $\tilde{\pi}(p, X)$ as given in our analysis of the second-stage problem. We estimate the parameters of the second-stage prior so as to fit as well as possible the empirically observed frequency distribution of lotteries (p, X) that reach the second stage. Thus the observed pattern of selection of the lotteries for which the second stage is reached is taken into account, but we have no need (for our purposes here) to estimate a model of the first-stage decision. That is left for future study.

F Stochastic Versions of Prospect Theory

In section 4.1 of the main text, we discuss the degree to which the bids of our average subject can be fit by a stochastic version of prospect theory. Note that a likelihood-based model comparison exercise of the kind that we undetake is possible only by augmenting prospect theory, as originally described by Kahneman and Tversky (1979) and Tversky and Kahneman (1992) with a model of random response errors, as in the empirical implementations of prospect theory reviewed by Stott (2006).

In all of the versions of prospect theory (PT) that we discuss, we assume that bids are drawn from the distribution (2.3), except that now f is a function of the objective data (p, X) rather than of a noisy internal representation. Here $f(p, X)$ is the logarithm of the

with either p or $|X|$ to any appreciable degree), while only one of these (subject 11) had similar difficulty making minimally sensible bids in the case of lotteries involving gains. (The peculiarity of the bids of these subjects is discussed further in the September 2022 draft of NBER Working Paper no. 30417, Appendix, section D.2.)

[112] Similarly, we assume two distinct information structures (internal representations), each with a separate information cost, for the two stages of the decision problem in Khaw *et al.* (2017).

[113] An empirical measure of this probability is given by one minus the fraction shown in Figure 8 for each lottery (p, X).

absolute value of the bid $\bar{C}$ implied by deterministic PT; thus we modify PT by multiplying the deterministic prediction $\bar{C}(p, X)$ by a log-normally distributed response error.[114] The deterministic prediction is the monetary amount $\bar{C}$ such that

$$V(\bar{C}; 1) \;=\; V(X; p), \tag{F.1}$$

where PT assigns a value (in non-monetary units) of

$$V(X; p) \;\equiv\; w(p) \cdot v(X)$$

to a random prospect offering the monetary amount X with probability p (and zero otherwise). Thus $\bar{C}$ is the amount of money such that, according to (deterministic) prospect theory, a DM should be indifferent between receiving $\bar{C}$ with certainty and receiving X with probability p.

We consider a variety of different specifications for the value function $v(X)$ and the probability weighting function $w(p)$, each of which has been popular in the empirical literature. In the *symmetric* versions of PT that we consider, we impose the restriction that $v(-X) = -v(X)$, in which case PT (like our noisy coding model) predicts that (except for the sign of the bids) the distribution of bids for any values of p and $|X|$ are the same in the case of both lotteries involving gains and lotteries involving losses. More generally, one might allow the function to be asymmetric. In the empirical fits reported in Tversky and Kahneman (1992), an asymmetric *power law* function is assumed: for values of X with either sign, it is assumed that

$$v(X) \;=\; \mathrm{sign}(X) \cdot |X|^{\alpha} \tag{F.2}$$

for some $0 < \alpha \leq 1$, with the value of α allowed to differ depending on the sign of X.[115] The cases called "power law" in Table 2 assume a *symmetric* power law function: a function of the form (F.2), with the same value of α regardless of the sign of X.[116]

The "power law" specification (F.2) has been very popular in empirical implementations of PT, as discussed by Stott (2006). But in the case of a power-law value function, equation (F.1) reduces to

$$\bar{C} \;=\; \left(\frac{w(p)}{w(1)}\right)^{\frac{1}{\alpha}} X,$$

which implies that the median value of WTP/EV (i.e., $\bar{C}/pX$) should be a function of p, independent of the value of X. Thus there should be no stake-size effects under this version

[114] As reviewed in Stott (2006), empirical implementations of PT often make the theory stochastic by *adding* a random term to $\bar{C}$ rather than assuming a multiplicative valuation error. However, the studies reviewed there generally model choices between pairs of lotteries, rather than elicited certainty equivalent values, as here. The use of a multiplicative response error specification makes the stochastic PT models that we consider more comparable to the variants of our cognitive noise model.

[115] Tversky and Kahneman also allow for a positive multiplicative factor different from 1, that can differ depending on the sign of X. (They argue for a larger multiplicative factor in the case of losses, reflecting loss aversion.) However, such a multiplicative factor has no consequences for the predicted valuations of prospects that involve payoffs that are all of the same sign, as in the case of the lotteries used in our experiment. Hence we can without loss of generality assume a multiplicative factor of 1, in the case of either gains or losses. We can instead separately identify the value of α for the cases of gains and losses respectively.

[116] We have also estimated variant models in which α is allowed to differ for gains and losses (results not reported here), but do not find an improvement in fit (increase of LL) sufficient to justify the additional free parameter, if the BIC is used to judge model fit. Hence we only report results for symmetric models here.

value function	prob. weighting	#params	α	γ	δ	ν_c
power law	linear	1	1			0.55
power law	TK92	2	1	0.540		0.23
power law	Prelec	4	0.849	0.519	0.838	0.21
logarithmic	TK92	3	0.001	0.542		0.23
logarithmic	Prelec	4	0.043	0.520	0.839	0.21

Table 7: Maximum-likelihood parameter estimates for the five stochastic versions of prospect theory referred to in Table 2 of the main text. The column "#params" indicates the number of free parameters penalized when computing the BIC statistics reported in Table 2.

of prospect theory. Other choices of value function can instead allow for stake-size effects, as discussed by Scholten and Read (2014). The most popular choice in the empirical literature focusing on stake-size effects has been the logarithmic specification advocated for example by Bouchouicha and Vieider (2017),

$$v(X) \;=\; \mathrm{sign}(X) \cdot \log(1 + \alpha X) \tag{F.3}$$

for some $\alpha > 0$. The cases called "logarithmic" in Table 2 assume a function of this kind with the same value of α for both gains and losses. We consider this specification in order to give PT as good a chance as possible to fit the stake-size effects that we find.

For the weighting function $w(p)$, the simplest case that we consider (called "linear" in Table 2) assumes that $w(p) = p$; in this case, PT is equivalent to a version of expected utility maximization, in which however the nonlinear utility function is applied to the *change* in wealth from an individual gamble rather than to the DM's overall wealth. (This version of expected utility theory does not correspond to the original proposal of Bernoulli, 1954 [1738], but is one that is commonly used in experimental studies of decision making under risk.) Tversky and Kahneman (1992) instead consider a one-parameter family of nonlinear weighting functions,

$$w(p) \;=\; \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}}, \tag{F.4}$$

for some $0 < \gamma \leq 1$. (Note that in the limiting case $\gamma = 1$, this reduces to the linear specification.[117] For values $0 < \gamma < 1$, the value function has an "inverse S shape" of the kind hypothesized by Kahneman and Tversky with limiting values $w(0) = 0$ and $w(1) = 1$.) As in the case of the value function, Tversky and Kahneman allow the parameter γ to be different for gains and losses; the "TK92" weighting function referred to in Table 2 in the main text is instead the symmetric case in which (F.4) holds with the same value of γ regardless of the sign of X.

A variety of other nonlinear probability weighting functions have been proposed in the literature, as reviewed by Stott (2006). Among these, the simple family (family of functions

[117] We nonetheless consider the linear case separately in Table 2, because assuming linearity *a priori* reduces the number of free parameters of the model. It would thus be possible for the model that imposes $w(p) = p$ to fit better, according to the BIC criterion, than the best-fitting member of the family of models that assume (F.4).

with two free parameters or fewer) that fits our data best is the two-parameter family proposed by Prelec (1998), in which

$$w(p) \ = \ \exp(-\delta(-\log p)^{\gamma}), \tag{F.5}$$

for parameter values $0 < \gamma, \delta \leq 1$. (This family also nests the linear specification in the case that $\gamma = \delta = 1$, while it implies an "inverse S shape" if γ and δ are both less than 1. Prelec derives this family of functions from an attractive set of axioms.) We present results for this alternative (called "Prelec" in Table 2) in order to show the case (among those that we have investigated) in which a stochastic version of PT is most successful in fitting our data.

The parameter values that best fit the data from our average subject are shown in Table 7, for each of several versions of PT. For each combination of functional forms for the value function and probability weighting function, the table shows the best-fitting parameter values when the same functions are used for both lotteries involving gains and those involving losses. In the case that the power law value function (F.2) is combined with either a linear probability weighting function (the expected utility case) or the TK92 weighting function, we find that the best-fitting parameter value, subject to the constraint that $\alpha \leq 1$, is $\alpha = 1$ (a linear value function). This means that the model on the first line of Table 7 is one in which both the value function and weighting function are linear; this corresponds to EV maximization, but with a (multiplicative) random response error. The single free parameter to estimate in this case is the value of ν_c. In the model on the second line, the constraint also binds, so that the predicted mean value of $\log(WTP/EV)$ for each value of p is due solely to the nonlinearity of the probability weighting function; the best-fitting value of γ is the one that best fits the implied function $w(p)$ to a graph of WTP/X as a function of p, as in the figures shown in Tversky and Kahneman (1992). This model has two free parameters to estimate: the weighting-function curvature parameter γ and the response noise parameter ν_c. When we instead allow the more flexible Prelec two-parameter family of weighting functions (F.5), the best-fitting value of α is somewhat less than 1, though the curvature of the best-fitting value function is still not severe.

If we instead assume the logarithmic family (F.3) of value functions, our conclusions about the best-fitting probability weighting functions are not much affected. (Compare the estimated values for γ on lines 2 and 4, or the estimated values for the Prelec parameters (γ, δ) on lines 3 and 5.) In fact, when we pair the logarithmic value function with the TK92 weighting function, the best-fitting value of α is near zero, meaning that the value function is estimated to be essentially linear (just as on line 2). When we instead pair this value function with the Prelec weighting function, the optimal value function again has more curvature; and the value function implied by the parameter value on line 5 is not shaped quite the same way as the one implied by the parameter value on line 3. The difference matters for the predicted stake-size effects; but it does not much affect the best-fitting parameter values for the probability weighting function. The consequences for overall model fit are shown in Table 2 in the main text.

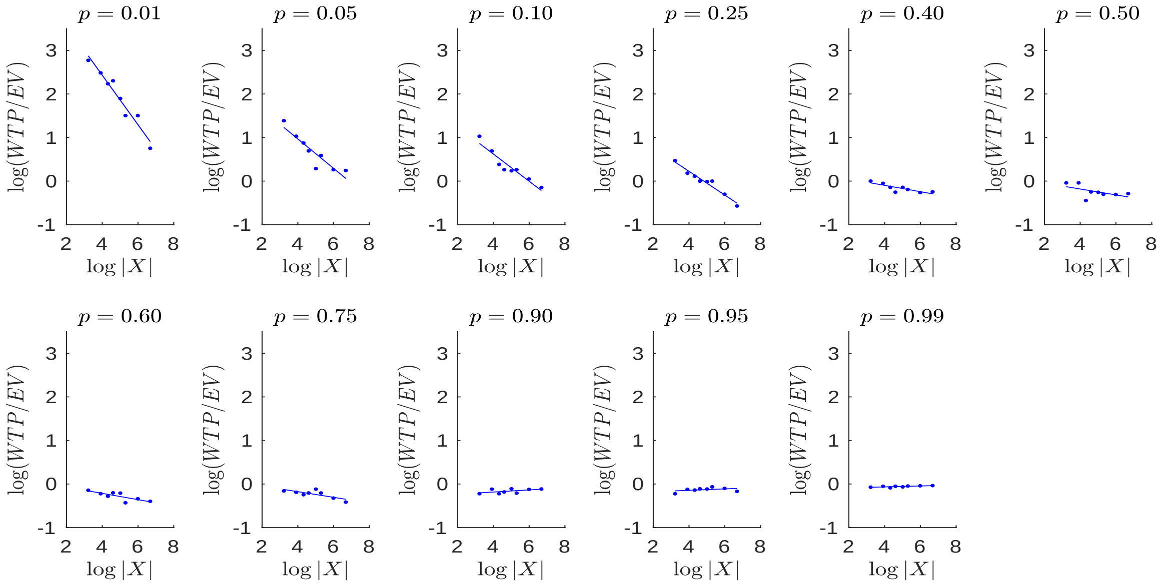

Figure 9: Fit of our baseline noisy coding model to the median bids in the study of Gonzalez and Wu (2022), treated as the bids of an average subject. The data (showing the bids) are the same as in Figure 9; the red lines show the predictions of the noisy coding model. The format is the same as in the top rows of Figures 5 and 6.

G Log-Linear Stake-Size Effects in the Data of Gonzalez and Wu (2022)

Among other studies of the valuation of simple lotteries, the study of Gonzalez and Wu (2022) is of particular interest for our purposes because, like us, they elicit certainty-equivalent values for lotteries that involve both a wide range of values of p and a wide range of monetary payoffs X. While their study also involves lotteries with more than one non-zero payoff, they consider a fairly large number of lotteries with only one non-zero payoff, like the ones used in our study; their results are directly comparable with ours on these trials. The lotteries of this kind that they use involve 8 different values of X (rather than only five, as in our study), and each of the 8 different values of X is paired with each of the 11 different values of p that they use. The probabilities that they consider also span a wider range, including values of p as small as 0.01 and as large as 0.99. The inclusion of a very small value of p is of particular interest, since we (like previous authors) find especially pronounced stake-size effects when p is small, and our theoretical model also implies that they should be especially extreme as p approaches zero.

Gonzalez and Wu (2022) also have a larger number of subjects than in our study: 47 subjects, each of whom is asked to value the same set of 165 different lotteries. There are however two disadvantages of their study relative to ours: First, they consider only lotteries involving potential gains, not lotteries involving potential losses as well. And second, they have each experimental subject value each lottery only once; thus they do not collect data on the amount of trial-to-trial variability in subjects' valuations of a given lottery.

It is nonetheless of interest to ask how the relative risk premia indicated by their data vary

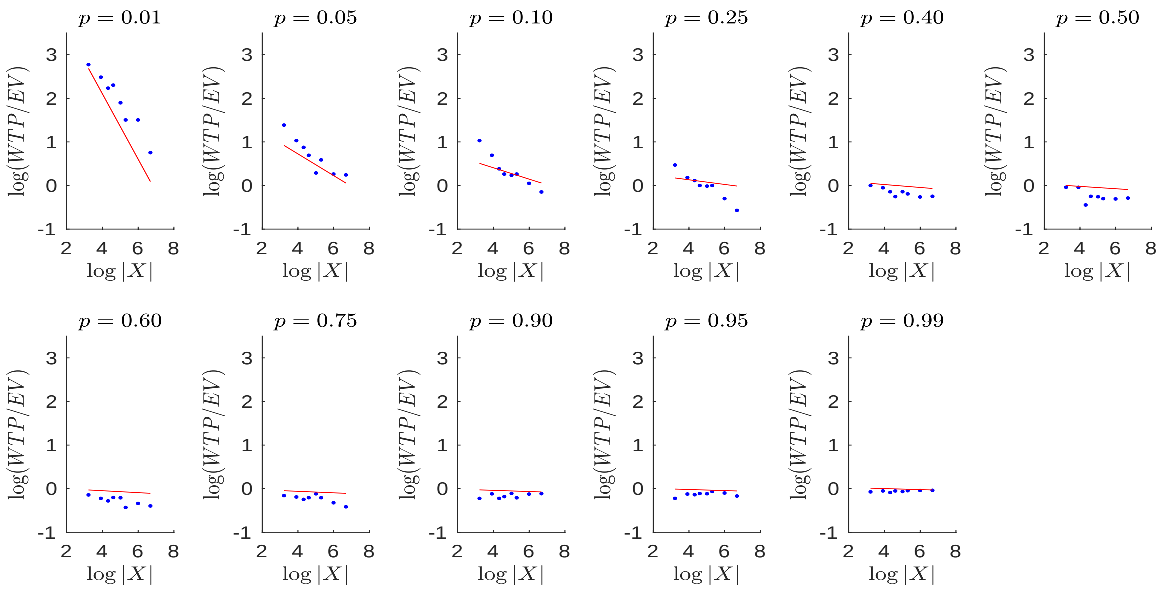

Figure 10: The value of WTP as a multiple of EV, for the median subject in the study of Gonzalez and Wu (2022). The format is the same as in the top rows of Figures 2 and 3 (here the data all refer to lotteries involving gains).

with p and X. We plot the median bid of their subjects for each lottery in Figure 9, using the same format as in the top row of Figures 2 and 3 of our paper.[118] As indicated by the linear regression lines included with the log-log plot in each panel, the slope of $\log(WTP/EV)$ as a function of $\log|X|$ is essentially zero for the highest values of p (all $p \geq 0.90$), but the relationship is downward-sloping for all lower values of p. Just as in our Figure 2, the most strongly negative-sloping relationships are observed for probabilities $p \leq 0.25$. The relationships are also approximately log-linear, as shown by the degree of fit of the linear regression lines.

Thus, to the extent that the data of Gonzalez and Wu (2022) can be used to address the same issues as our data, they confirm the regularities that we have noted in section 1 of the main text.[119] Indeed, they provide even stronger evidence for the effects that we document, in two important respects. First, Gonzalez and Wu consider a wider range of values for the stake size $|X|$: their largest monetary payoff (800) is 32 times as large as the smallest (25), whereas our largest payoff is only 4 times the size of our smallest; it is thus even more notable that WTP/EV appears to be a log-linear function of stake size in their data. And second, they consider a much smaller value of p (namely, 0.01) than our smallest value (0.05). They find that WTP/EV has a more strongly negative elasticity with respect to stake size when $p = 0.01$ than when $p = 0.05$ or 0.10; this is an even stronger confirmation of our conclusion that the slope becomes particularly negative for low values of p.

[118] In Figure 9, there are no intervals around the median bids shown, because we have no data on trial-to-trial variation, which is what the whiskers in Figures 2 and 3 indicate.

[119] This is reassuring, especially in light of the many differences in their experimental procedure: the values of X that they use are all round numbers, their subjects are not required to value as large a number of lotteries over the course of the session as ours are, certainty-equivalents are elicited using a multiple-price list in their case, etc.

Cognitive Noise Models			
model		LL	BIC
baseline model		-320.6	663.5
Stochastic Prospect Theory			
value function	prob. weighting		
power law	linear	-408.5	830.3
power law	TK	-334.2	690.7
power law	Prelec	-323.8	674.3
logarithmic	TK	-328.5	674.8
logarithmic	Prelec	-316.6	655.5

Table 8: Model comparison statistics for the fit of several stochastic models of lottery valuation to the median bids reported by Gonzalez and Wu (2022), treated as the bids of an average subject.

We can also test the fit of our baseline model to the bids of the average subject of Gonzalez and Wu. Here we treat the single value of $y_j \equiv \log(WTP_j/EV_j)$ elicited (from the median subject) for each lottery (p_j, X_j) as a single draw from the predicted distribution $N(m_j, v_j)$, where m_j and v_j depend on the values of p_j and X_j (and model parameters) in the same way as has been explained above. For any hypothesized model parameters, the log likelihood of the data is then given by $LL = \sum_j L_j$, where for each of the 88 single-nonzero-outcome lotteries j, we can again write L_j as the sum of two terms, as in (D.3). Here the expression $L_1(p_j, X_j)$ is again defined as in (D.4), while L_{2j} is now defined more simply as

$$L_{2j} \;=\; -\frac{1}{2v_j}(y_j - m_j)^2 \;-\; \frac{1}{2}\log(2\pi v_j),$$

instead of as in (D.7). We can then estimate our model parameters so as to maximize LL.

The fit of the resulting model predictions to the data plotted in Figure 9 is displayed visually in Figure 10. The dots represent the same data (the bids of the average subject) as in Figure 9, but instead of the atheoretical linear regression lines of the earlier figure, the red lines in Figure 10 plot the theoretical predictions $y_j = \alpha_p + \beta_p \log X_j$ in each panel (corresponding to a different value of p). We find that our theoretical model is broadly consistent with the data of Gonzalez and Wu (2022) as well, though the best-fitting parameter values are different than in the case of our subjects.[120]

The log likelihood of the data when the parameters of the cognitive noise model are optimized is indicated in Table 8. For purpose of comparison, the table also shows the

[120]The maximum likelihood parameter estimates using the Gonzalez-Wu data are $A = 0.0004$, $\nu_z = 0.84$, $\nu_c = 0.000037$. Thus all three noise parameters are smaller when the model is fit to the bids of the average subject of Gonzalez and Wu. At least part of the difference probably reflects the fact that Gonzalez and Wu consider only lotteries involving gains; also in the case of our subjects, if we fit the model separately to the bids on lotteries involving only gains, we obtain somewhat smaller noise parameters than the ones reported in Table 4 for our baseline model: $A = 0.0008$, $\nu_z = 1.50$, and $\nu_c = 0.28$. Some of the difference may also reflect the fact that the subjects of Gonzalez and Wu do not have to value as large a number of lotteries, and thus may be less affected by time pressure or fatigue; recall that in Table 4 we also obtain smaller noise parameter estimates in the case of the subjects who value only 400 lotteries.

corresponding values for the log likelihood LL and the BIC statistic for five stochastic variants of prospect theory (the same five as are compared to our model in Table 2 of the main text). We see that the fit of our model to the data of Gonzalez and Wu (2022) is better than that of a number of common quantitative specifications of prospect theory, though there is at least one version of prospect theory that fits somewhat better than our baseline cognitive noise model: this is a model that combines the logarithmic value function of Bouchouicha and Vieieder (2017) with the probability weighting function proposed by Prelec (1998), and adds a log-normal multiplicative response error to the DM's bid. It is also worth noting that the best-fitting version of prospect assumes much larger random response errors ($\nu_c = 0.058$) than does our baseline model. This is because in the case of prospect theory, any failure of the median bid to precisely fit the value implied by deterministic prospect theory must be attributed to response error; in the noisy coding model, instead, responses would be predicted to be random even in the absence of response errors (i.e., if we set ν_c equal to zero).

References

[1] Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer, "Time Will Tell: Recovering Preferences When Choices Are Noisy," *Journal of Political Economy* 129: 1828-1877 (2021).

[2] Augenblick, Ned, Eben Lazarus, and Michael Thaler, "Overinference from Weak Signals and Underinference from Strong Signals," MIT Sloan Research Paper no. 6830-22, revised May 2023.

[3] Azeredo da Silveira, Rava, Yeji Sung, and Michael Woodford, "Optimally Imprecise Memory and Biased Forecasts," NBER Working Paper no. 28075, November 2020.

[4] Barretto-García, Miguel, Gilles de Hollander, Marcus Grueschow, Rafael Polania, Michael Woodford, and Christian C. Ruff, "Individual Risk Attitudes Arise from Noise in Neurocognitive Magnitude Representations," *Nature Human Behaviour* 7: 1551-1567 (2023).

[5] Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, "Measuring Utility By a Single-Response Sequential Method," *Behavioral Science* 9: 226-232 (1964).

[6] Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro, "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Journal of the European Economic Association* 11: 1231-1255 (2013).

[7] Bhui, Rahul, and Yang Xiang, "A Rational Account of the Repulsion Effect," *PsyArXiv* preprint, posted November 9, 2022.

[8] Bouchouicha, Ranoua, and Ferdinand M. Vieider, "Accommodating Stake Effects Under Prospect Theory," *Journal of Risk and Uncertainty* 55: 1-28 (2017).

[9] Burnham, Kenneth P., and Anderson, David R., *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,* 2d. ed., New York: Springer, 2002.

[10] Charles, Constantin, Cary Frydman, and Mete Kilic, "Insensitive Investors," *Journal of Finance,* forthcoming.

[11] Charness, Gary, Nir Chemaya, and Dario Trujano-Ochoa, "Learning Your Own Risk Preferences," *Journal of Risk and Uncertainty* 67: 1-19 (2023).

[12] Choi, Syngjoo, Jeongbin Kim, Eungik Lee, and Jungmin Lee, "Probability Weighting and Cognitive Ability," *Management Science* 68: 5201-5215 (2022).

[13] Daunizeau, Jean, "Semi-Analytical Approximations to Statistical Moments of Sigmoid and Softmax Mappings of Normal Variables," preprint posted on *arXiv,* March 3, 2017.

[14] Deck, Cary, and Salar Jahedi, "The Effect of Cognitive Load on Economic Decision Making: A Survey and New Experiments," *European Economic Review* 78: 97-119 (2015).

[15] Dehaene, Stanislas, *The Number Sense*, revised and updated edition, Oxford: Oxford University Press, 2011.

[16] Dehaene, Stanislas, Emmanuel Dupoux, and Jacques Mehler, “Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison,” *J. Experimental Psychology: Human Perception and Performance* 16: 626-641 (1990).

[17] Dehaene, Stanislas, and J. Frederico Marques, “Cognitive Euroscience: Scalar Variability in Price Estimation and the Cognitive Consequences of Switching to the Euro,” *Quarterly Journal of Experimental Psychology A* 55: 705-731 (2002).

[18] Eckert, Johanna, Josep Call, Jonas Hermes, Esther Herrmann, and Hannes Rakoczy, “Intuitive Statistical Inferences in Chimpanzees and Humans Follow Weber’s Law,” *Cognition* 180: 99-107 (2018).

[19] Enke, Benjamin, and Thomas Graeber, “Cognitive Uncertainty in Intertemporal Choice,” NBER Working Paper no. 29577, December 2021.

[20] Enke, Benjamin, and Thomas Graeber, “Cognitive Uncertainty,” working paper, Harvard University, January 2023.

[21] Enke, Benjamin, Thomas Graeber, and Ryan Oprea, “Complexity and Time,” NBER Working Paper no. 31047, March 2023.

[22] Enke, Benjamin, and Cassidy Shubatt, “Quantifying Lottery Choice Complexity,” working paper, Harvard University, June 2023.

[23] Ert, Eyal, and Ernan Haruvy, “Revisiting Risk Aversion: Can Risk Preferences Change with Experience?” *Economics Letters* 151(C): 91-95 (2017).

[24] Fechner, Gustav, *Elements of Psychophysics,* New York: Holt, Rinehart and Winston, 1966. (English translation of the German original, published 1860.)

[25] Fehr-Duda, Helga, Adrian Bruhin, Thomas Epper, and Renate Schubert, “ Rationality on the Rise: Why Relative Risk Aversion Increases with Stake Size,” *Journal of Risk and Uncertainty* 40: 147-180 (2010).

[26] Findling, Charles, and Valentin Wyart, “Computation Noise in Human Learning and Decision-Making: Origin, Impact, Function,” *Current Opinion in Behavioral Sciences* 38: 124-132 (2021).

[27] Frydman, Cary, and Lawrence J. Jin, “Efficient Coding and Risky Choice,” *Quarterly Journal of Economics* 137: 161-213 (2022).

[28] Frydman, Cary, and Lawrence J. Jin, “On the Source and Instability of Probability Weighting,” working paper, U.S.C., February 2023.

[29] Gabaix, Xavier, and David Laibson, “Myopia and Discounting,” NBER Working Paper no. 23254, revised May 2022.

[30] Gerhardt, Holger, Guido P. Biele, Hauke R. Heekeren, and Harald Uhlig, "Cognitive Load Increases Risk Aversion," SFB 649 Discussion Paper no. 2016-011, Humboldt University Berlin, March 2016.

[31] Gershman, Samuel J., and Rahul Bhui, "Rationally Inattentive Intertemporal Choice," *Nature Communications* 11: 3365 (2020).

[32] Gold, Joshua I., and Hauke R. Heekeren, "Chapter 19: Neural Mechanisms for Perceptual Decision Making," in P.W. Glimcher and E. Fehr, eds., *Neuroeconomics: Decision Making and the Brain,* 2d edition, London: Academic Press, 2014.

[33] Gonzalez, Richard, and George Wu, "On the Shape of the Probability Weighting Function," *Cognitive Psychology* 38: 129-166 (1999).

[34] Gonzalez, Richard, and George Wu, "Composition Rules in Original and Cumulative Prospect Theory," *Theory and Decision* 92: 647-675 (2022).

[35] Green, David M., and John A. Swets, *Signal Detection Theory and Psychophysics,* New York: Wiley, 1966.

[36] Heng, Joseph A., Michael Woodford, and Rafael Polania, "Efficient Numerosity Estimation Under Limited Time," preprint posted on *bioRXiv,* July 3, 2023.

[37] Hershey, John C., and Paul J. Schoemaker, "Prospect Theory's Reflection Hypothesis: A Critical Examination," *Organizational Behavior and Human Performance* 25: 395-418 (1980).

[38] Hertwig, Ralph, and Ido Erev, "The Description-Experience Gap in Risky Choice," *Trends in Cognitive Sciences* 13: 517-523 (2009).

[39] Hertwig, Ralph, Greg Barron, Elke U. Weber, and Ido Erev, "Decisions from Experience and the Effect of Rare Events in Risky Choice," *Psychological Science* 15: 534-539 (2004).

[40] Hollands, J.G., and Brian P. Dyre, "Bias in Proportion Judgments: The Cyclical Power Model," *Psychological Review* 107: 500-524 (2000).

[41] Holt, Charles A., and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review* 92: 1644-1655 (2002).

[42] Holt, Charles A., and Susan K. Laury, "Risk Aversion and Incentive Effects: New Data without Order Effects," *American Economic Review* 95: 902-904 (2005).

[43] Kachelmeier, Steven J., and Mohamed Shehata, "Examining Risk Preferences under High Monetary Incentives: Experimental Evidence from the People's Republic of China," *American Economic Review* 82: 1120-1141 (1992).

[44] Kahneman, Daniel, "Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice," Nobel Prize Lecture, December 8, 2002.

[45] Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* 47: 263-291 (1979).

[46] Khaw, Mel W., Ziang Li, and Michael Woodford, "Cognitive Imprecision and Small-Stakes Risk Aversion," *Review of Economic Studies* 88: 1979-2013 (2021).

[47] Khaw, Mel W., Luminita L. Stevens, and Michael Woodford, "Discrete Adjustment to a Changing Environment: Experimental Evidence," *Journal of Monetary Economics* 91: 88-103 (2017).

[48] Kirchler, Michael, David Andersson, Caroline Bonn, Magnus Johannesson, Erik Ø. Sørenson, Matthias Stefan, Gustav Tinghög, and Daniel Västfjäll, "The Effect of Fast and Slow Decisions on Risk Taking," *Journal of Risk and Uncertainty* 54: 37-59 (2017).

[49] Krueger, Lester E., "Perceived Numerosity: A Comparison of Magnitude Production, Magnitude Estimation, and Discrimination Judgments," *Perception and Psychophysics* 35: 536-542 (1984).

[50] Lucas, Robert E., Jr., "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1: 19-46 (1976).

[51] Markowitz, Harry, "The Utility of Wealth," *Journal of Political Economy* 60: 151-158 (1952).

[52] Mosteller, Frederick, and Philip Nogee, "An Experimental Measurement of Utility," *Journal of Political Economy* 59: 371-404 (1951).

[53] Moyer, Robert S., and Thomas K. Landauer, "Time Required for Judgments of Numerical Inequality," *Nature* 215: 1519-1520 (1967).

[54] Myatt, David P., and Chris Wallace, "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies* 79: 340-374 (2012).

[55] Natenzon, Paulo, "Random Choice and Learning," *Journal of Political Economy* 127: 419-457 (2019).

[56] Netzer, Nick, Arthur Robson, Jakub Steiner, and Pavel Kocourek, "Endogenous Risk Attitudes," working paper, University of Zurich, August 2022.

[57] Oprea, Ryan, "Simplicity Equivalents," working paper, UC Santa Barbara, June 2023.

[58] Petzschner, Frederike H., Stefan Glasauer, and Klaas E. Stephan, "A Bayesian Perspective on Magnitude Estimation," *Trends in Cognitive Sciences* 19: 285-293 (2015).

[59] Phillips, Lawrence D., and Ward Edwards, "Conservatism in a Simple Probability Inference Task," *Journal of Experimental Psychology* 72: 346-354 (1966).

[60] Plott, Charles R., "Rational Individual Behavior in Markets and Social Choice Processes: The Discovered Preference Hypothesis," in K.J. Arrow, E. Colombatto, M. Perleman, and C. Schmidt, eds., *Rational Foundations of Economic Behavior*, London: Macmillan, 1996.

[61] Porcelli, Anthony J., and Mauricio R. Delgado, "Acute Stress Modulates Risk Taking in Financial Decision Making," *Psychological Science* 20: 278-283 (2009).

[62] Prelec, Drazen, "The Probability Weighting Function," *Econometrica* 66: 497-527 (1998).

[63] Rubinstein, Ariel, "Response Time and Decision Making: An Experimental Study," *Judgment and Decision Making* 8(5): 540-551 (2013).

[64] Sargent, Thomas J., "Lucas' Critique," in *Macroeconomic Theory,* 2d ed., Orlando: Academic Press, 1987, pp. 397-398.

[65] Scholten, Mark, and Daniel Read, "Prospect Theory and the 'Forgotten' Fourfold Pattern of Risk Preferences," *Journal of Risk and Uncertainty* 48: 67-83 (2014).

[66] Steiner, Jakub, and Colin Stewart, "Perceiving Prospects Properly," *American Economic Review* 106: 1601-1631 (2016).

[67] Stott, Henry P., "Cumulative Prospect Theory's Functional Menagerie," *Journal of Risk and Uncertainty* 32: 101-130 (2006).

[68] Thurstone, Leon L., *The Measurement of Value,* Chicago: University of Chicago Press, 1959.

[69] Tversky, Amos, and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5: 297-323 (1992).

[70] Van de Kuilen, Gijs, "Subjective Probability Weighting and the Discovered Preference Hypothesis," *Theory and Decision* 67: 1-22 (2009).

[71] Van de Kuilen, Gijs, and Peter P. Wakker, "Learning in the Allais Paradox," *Journal of Risk and Uncertainty* 33: 155-164 (2006).

[72] Van den Berg, Ronald, and Wei Ji Ma, "A Resource-Rational Theory of Set-Size Effects in Human Visual Working Memory," *eLife* 7:e34963 (2018).

[73] Vieider, Ferdinand M., "Noisy Coding of Time and Reward Discounting," working paper, Ghent University, December 2021.

[74] Vieider, Ferdinand M., "Decisions Under Uncertainty as Bayesian Inference on Choice Options," *Management Science,* forthcoming.

[75] Wei, Xue-Xin, and Alan A. Stocker, "A Bayesian Observer Model Constrained by Efficient Coding Can Explain 'Anti-Bayesian' Percepts," *Nature Neuroscience* 18: 1509-1517 (2015).

[76] Wei, Xue-Xin, and Alan A. Stocker, "Lawful Relation Between Perceptual Bias and Discriminability," *Proceedings of the National Academy of Sciences* 114: 10244-10249 (2017).

[77] Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion *et al, Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps,* Princeton University Press, 2003.

[78] Woodford, Michael, "Prospect Theory as Efficient Perceptual Distortion," *American Economic Review* 102(3): 1-8 (2012).

[79] Woodford, Michael, "Individualistic Welfare Analysis in the Age of Behavioral Science," *Capitalism and Society,* vol. 13, issue 1, article 3, posted online October 15, 2018.

[80] Woodford, Michael, "Modeling Imprecision in Perception, Valuation and Choice," *Annual Review of Economics* 12: 579-601 (2020).

[81] Young, Diana L., Adam S. Goodie, Daniel B. Hall, and Eric Wu, "Decision Making Under Time Pressure, Modeled in a Prospect Theory Framework," *Organizational Behavior and Human Decision Processes* 118: 179-188 (2012).

[82] 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, article 1 (2012).

[83] Zhang, Hang, Xiangjuan Ren, and Laurence T. Maloney, "The Bounded Rationality of Probability Distortion," *Proceedings of the National Academy of Sciences* 117: 22024-22034 (2020).