CESifo WORKING PAPERS

9923
2022
September 2022

Cognitive Imprecision and Stake-Dependent Risk Attitudes

Mel Win Khaw, Ziang Li, Michael Woodford

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9923*

Cognitive Imprecision and Stake-Dependent Risk Attitudes

Abstract

In an experiment that elicits subjects' willingness to pay (WTP) for the outcome of a lottery, we confirm the fourfold pattern of risk attitudes described by Kahneman and Tversky. In addition, we document a systematic effect of stake sizes on the magnitude and sign of the relative risk premium, holding fixed both the probability that a lottery pays off and the sign of its payoff (gain vs. loss). We further show that in our data, 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 behaviour 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.

JEL-Codes: C910, D030, D810, D870.

Keywords: prospect theory, fourfold pattern, endogenous precision, cognitive noise.

Mel Win Khaw
Microsoft Corporation
melwinkhaw@gmail.com

Ziang Li
Princeton University / NJ / USA
ziang.li@princeton.edu

Michael Woodford
Columbia University / New York / NY / USA
michael.woodford@columbia.edu

August 23, 2022
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 Helga Fehr-Duda and Peter Wakker 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. In particular, the celebrated "fourfold pattern" of risk attitudes documented by Kahneman and Tversky (1979; also Tversky and Kahneman, 1992) asserts that people tend to be risk-averse when evaluating simple gambles involving gains when the probability of the larger gain is large, or when evaluating one involving losses when the probability of the larger loss is small, but that the same people tend to risk-seeking when evaluating random gains when when the probability of the larger gain is small, or random losses when the probability of the larger loss is large.

Prospect theory interprets this fourfold pattern as resulting from the inverse-S shape of the probability weighting function. Standard expositions (following Tversky and Kahneman, 1992) suppose that the shape of the probability weighting function can be inferred by observing how the ratio of a lottery's certainty-equivalent value CE to its expected value EV varies with the probability p of the larger payoff. Thus Tversky and Kahneman (1992) plot the median value of CE/EV for a variety of lotteries corresponding to different values of p, with separate plots for lotteries involving random gains and random losses respectively.[1] While they plot separate values of these ratios for lotteries involving larger and smaller (hypothetical) monetary amounts, Tversky and Kahneman conclude that the ratio CE/EV for a given value of p does not depend much on the stake size, justifying their emphasis on how the relative risk premium varies with p, and their intepretation of the plots as showing the shape of the probability weighting function.

However, a considerable body of evidence suggests that there are systematic effects of stake size on apparent risk attitudes. It has often been remarked that relative risk aversion appears to be greater when stakes are larger — at least when the payoffs are real (as opposed to hypothetical), and large enough to matter to the subjects (Holt and Laury, 2002, 2005). This finding generally relates to the evaluation of random gains with a relatively large probability of the larger possible gain (only one of the cells of Kahneman and Tversky's fourfold pattern). Another branch of the literature (Hershey and Schoemaker, 1980; Scholten and Read, 2014), however, considers gambles in which there is only a small probability of a non-zero gain, and finds that while choices are risk-seeking in the case of small enough potential gains (in accordance with Kahneman and Tversky's fourfold pattern), they are instead risk-averse for the same subjects when the potential gains are large enough. Similarly, when there is only a small probability of a non-zero loss, these papers find that choices are risk-averse in the case of small enough potential losses, but become risk-seeking when potential losses are large enough.[2] The latter finding indicates that stake effects cannot simply be attributed to greater caution when more is at stake.

Fehr-Duda *et al.* (2010) study the effects of stake size more systematically, eliciting certainty equivalents for a range of lotteries in both the gain and loss domains, and for a wide range of values of p, as well as several different stake sizes for each value of p. Their results regarding the sign of the relative risk premium (i.e., whether CE/EV is greater or less

[1] See their Figures 1 and 2.

[2] Scholten and Read (2014) call this an alternative ("forgotten") fourfold pattern, one that depends on the size of the stakes as opposed to the size of the probability of a non-zero payoff, and note that it had been conjectured as early as in the work of Markowitz (1952).

than 1) conform to Kahneman and Tversky's fourfold pattern;[3] however, they find nontrivial stake effects on the *size* of the relative risk premium, especially in the gain domain. In the case of random gains, the pattern is clear: for any value of p, the median ratio CE/EV is *smaller* when stakes are larger.[4] (For large p, this means greater risk aversion, while in the case of small p, it means less extreme risk-seeking.) While the qualitative effect is the same for all values of p, it is largest when p is small; this may explain why particularly large stake-size effects have been found in other experiments involving very small probabilities of a non-zero payoff.[5] A similar effect of stake size on the median ratio CE/EV is found in the loss domain, though less consistently.

None of these stake-size effects are consistent with prospect theory, at least if (as is common in quantitative tests of the theory), one assumes an isoelastic value function. Generalizations of prospect theory have been proposed to allow for stake-size effects, but these often have other unappealing implications,[6] and one may wonder why preferences over even very simple lotteries should be so complex. Here we propose a novel explanation for the pattern of stake-size effects just summarized, that does not depend on assuming that the objective served by subjects' decision rules is complex at all.

In previous work (Khaw *et al.,* 2021), we propose 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; the decision rules are posited to be optimal in the sense of maximizing the expected financial wealth of the decision maker (DM).[7] In the earlier paper, this idea was illustrated in the context of an incentivized experiment in which subjects chose between a small random gain and a small certain amount; because the probability p of the non-zero gain was the same on all trials, the cognitive noise that was emphasized in that paper was noise in the internal representations of the two monetary payoffs that defined the decision problem on a given trial: the potential gain from the risky lottery and the alternative certain gain.[8]

Here we apply the same idea to an experiment of the kind in Tversky and Kahneman (1992), or in Fehr-Duda *et al.* (2010), in which certainty-equivalent values are solicited for simple lotteries, that involve either random gains or random losses, and a wide range of

[3]Here it should be noted that the range of variation in stake sizes that they consider is much smaller than those in the experiments of Hershey and Schoemaker (1980) or Scholten and Read (2014). Thus the results of these authors are not inconsistent with those obtained by Fehr-Duda *et al.* (2010) for small values of p.

[4]See in particular the upper right panel of their Figure 2.

[5]Kachelmeier and Shehata (1992), one of the earliest studies to report a stake-size effect, also found greater stake-sensitivity in the case of small probabilities.

[6]Thus Scholten and Read (2014) conclude that "stake dependence emerges as a major challenge to prospect theory, and to other theories of choice under risk" (p. 82).

[7]Our argument is not that people are risk-neutral even in the case of large gambles, but rather that there should be little change in their marginal utility of additional money income in the case of different outcomes of a small-stakes laboratory experiment. See further discussion in our earlier paper.

[8]Frydman and Jin (2022) and Garcia *et al.* (2022) provide additional support for this model. The proposal of Khaw *et al.* (2021) 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, Bhui and Xiang (2022), Enke and Graeber (2021, 2022), Frydman and Nunnari (2022), Gabaix and Laibson (2017), Natenzon (2019), Steiner and Stewart (2016), Thaler (2021), and Woodford (2012, 2020).

different values of p on different trials.[9] Because p varies from trial to trial in the experiment considered here (in addition to variation in the monetary payoff offered by the lottery), it is reasonable to expect that this information will also be encoded and/or retrieved with noise, and the introduction of a noisy internal representation of p is an important innovation in the model presented here. We also now allow the precision with which the potential monetary payoff is encoded to vary depending on the probability of its being received; this feature of the model presented here is also one that did not arise in the simpler decision problem considered in our earlier paper. (We allow for variable precision, without a proliferation of additional free parameters, by supposing that the precision of encoding for each perceived probability of the non-zero outcome is optimized given a linear cost of greater precision; the unit cost of precision is then the single free parameter determining the precision of encoding of monetary amounts.) Finally, because subjects must supply an estimate of the certainty equivalent in the current experiment, from among a continuum of possible responses rather than simply making a binary choice, the current problem also requires us to model imprecision in response selection. Since the response is expressed in monetary units, our model of imprecision in response selection embodies the same logarithmic model of imprecision in mental manipulations of numerical magnitudes as is assumed in our model of the encoding of the lottery payoff (introduced and motivated in Khaw *et al.*, 2021).

While the application to the current experiment involves several new complications, the model remains a fairly parsimonious one. Apart from the specification of the prior distribution over lotteries for which the DM's decision rule is assumed to be optimized (parameters determined by the range of lotteries used in the experiment), there are only three free parameters that determine the model's quantitative predictions: one specifying the imprecision with which probabilities are encoded, another specifying the imprecision of the encoding of the monetary payoffs, and a third specifying the imprecision of response selection. This is no more free parameters than even the most parsimonious empirical implementations of prospect theory require;[10] moreover, our three parameters are conceptually related, as all of them simply specify the degree of noise in cognitive operations involved in the DM's task. The model nonetheless does a fairly good job of simultaneously explaining both the average valuations of a given lottery and the degree of variability of the valuations, as a function of the lottery characteristics, in both the gain and loss domains, for a range of sizes of the probability p of the non-zero payoff, and for a range of stake sizes.

In particular, our model successfully explains stake-size effects of the kind discussed above (and that we document in our data). The model predicts that in both the gain and loss domains, and regardless of the size of p, the median value of CE/EV should be a decreasing function of the stake size. It further makes a much more specific prediction: that for each value of p, $\log(CE/EV)$ should be an affine function of the logarithm of the absolute value

[9] Enke and Graeber (2022) also attribute biases in the valuations elicited in experiments of this kind to cognitive imprecision, though modeled somewhat differently than we do here. Their slightly different concept of cognitive imprecision is discussed further in section 5.1 below. Unlike the current study, Enke and Graeber (2022) do not measure or seek to explain stake-size effects.

[10] At a minimum, empirical versions of prospect theory involve parameters specifying the degree of curvature of the value function, the degree of curvature of the probability-weighting function, and the degree of noisiness of the DM's choices (Stott, 2006). Often, the number of parameters is larger; for example, separate parameters are fit to choices involving random losses rather than random gains.

of the monetary payoff, with a slope between - 1 and 0 (that can depend on p); and we show that this prediction is fairly well satisfied in our data. It implies that the relationship between median CE/EV and the stake size should be identical in the gain and loss domains, and this is also nearly true in our data. Finally, it implies that the degree of stake-sensitivity should be greater the lower the precision of the encoding of the monetary payoff, and that this should be lower the lower p is perceived to be. Hence the model predicts both that valuations should be noisier (in percentage terms) and that stake-sensitivity should be greater, when the probability p is smaller; and both predictions are verified in our data.

In section 1, we present the results of our new experiment, emphasizing the pattern of stake-dependence of risk attitudes that we observe. Section 2 then presents and motivates the elements of our theoretical model, with particular attention to the novel elements relative to the model presented in Khaw *et al.* (2021). Section 3 derives and tests a first set of predictions of the model, that are independent of the precise way in which we model the noisy encoding of probabilities or the endogenous variation in the precision of the encoding of monetary payoffs. This section explains both why the relative risk premium is predicted to be stake-dependent, and why the degree of stake-sensitivity should covary with the degree of imprecision of encoding of monetary payoffs. Section 4 then derives the further predictions of our model of noisy internal representation of probabilities and endogenous imprecision, and compares the complete predictions of our three-parameter model to our dataset. Section 5 discusses how the model can be further extended to address the findings of certain related experiments, and section 6 concludes.

1 Stake-Dependent Risk Attitudes: New Experimental Evidence

Here we provide additional evidence regarding the stake-dependence of risk attitudes through a new experimental study. As in the previous studies of Tversky and Kahneman (1992) and Fehr-Duda *et al.* (2010), 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. We do more than simply replicate the results of these authors, 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 CE 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 subjects' valuations than in the earlier studies, also in order to make the degree of random trial-by-trial variation more visible.[11]

[11] 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 is demonstrated by Holt and Laury (2002, 2005).

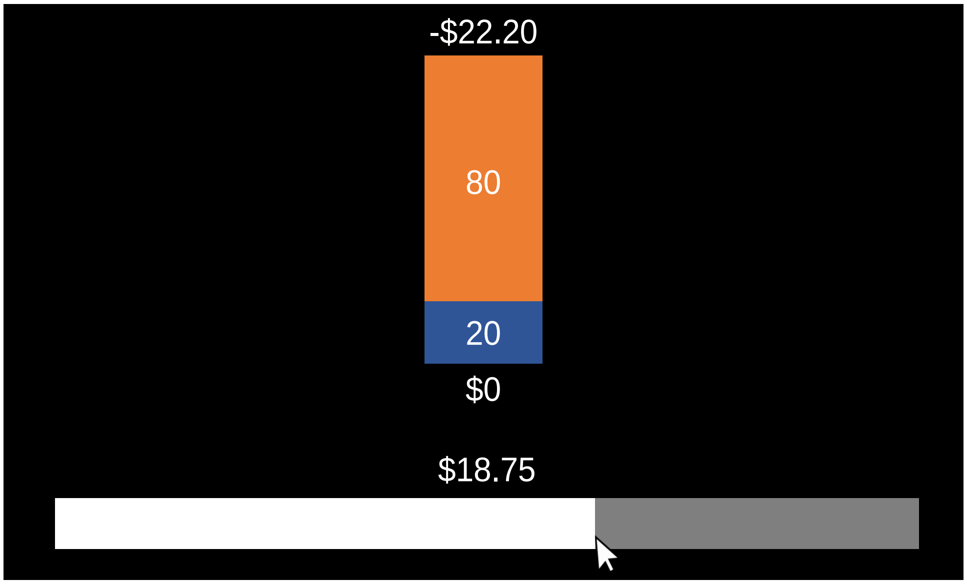

Figure 1: Example of the screen seen by an experimental subject.

1.1 Experimental Design

A total of 24 subjects[12] 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.[13] Five different values of the non-zero payoff

[12]These were student subjects recruited at Columbia University, following procedures approved by the Columbia Institutional Review Board under protocol IRB-AAAQ2255. Four other subjects also performed the experiment, but their data have been excluded on the ground that these subjects appeared not sufficiently engaged with the experimental task. The grounds for exclusion, and some ways in which the data of the excluded subjects compare with those of the other subjects, are discussed in the Appendix, section D.2.

[13]The full set of 11 different probabilities were not used with any of the subjects. Instead, 12 of the subjects completed 400 trials each, in which five values of p were used; the other 12 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 D.1.

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.[14] 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).[15]

On some trials, subjects submit a bid of zero (the leftmost position of the slider).[16] 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. (We discuss our theoretical interpretation of the zero-bid trials further in section 2.6 below.)

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 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 $\log(WTP/EV)$,

[14] 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.

[15] The incentives created by this procedure are discussed further below, in section 2.4.

[16] This occurs about 1.4 percent of the time overall, though more frequently when the EV of the lottery is small. See the Appendix, section D.3, for more information about these bids.

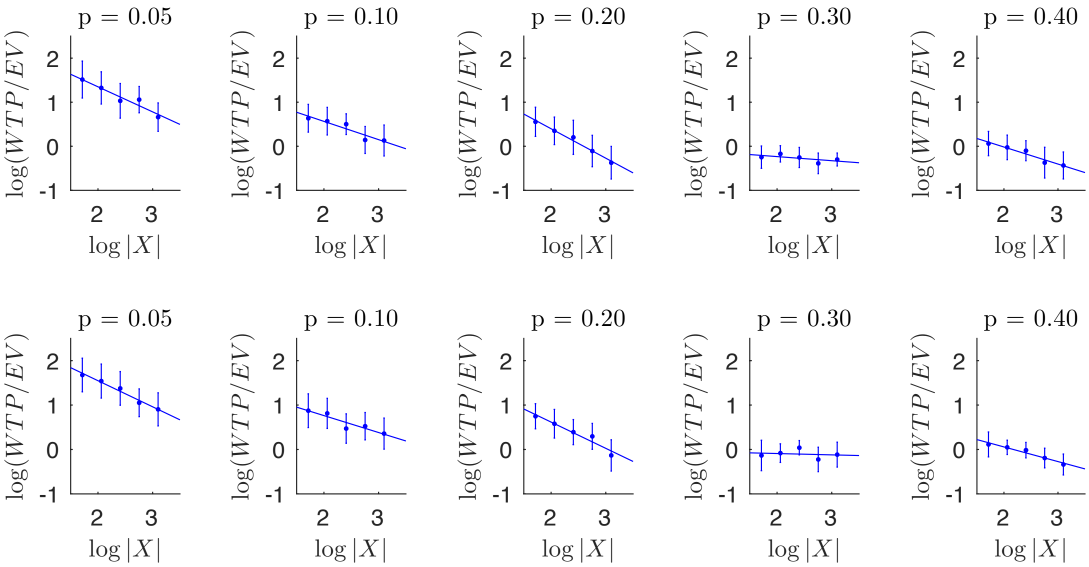

Figure 2: The distribution of values for WTP as a multiple of EV, for lotteries with different values of p (the different columns) and $|X|$ (the horizontal axis within each panel). The top panel in each column refers to lotteries involving random gains ($X > 0$), and the bottom panel to lotteries involving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

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 reward function explained in equation (2.7) 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.[17]

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.[18] The horizontal line in each panel indicates the prediction of an OLS regression model (with separate coefficients for each panel). Figure 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

[17]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.

[18]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 4 and 5). 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 C.

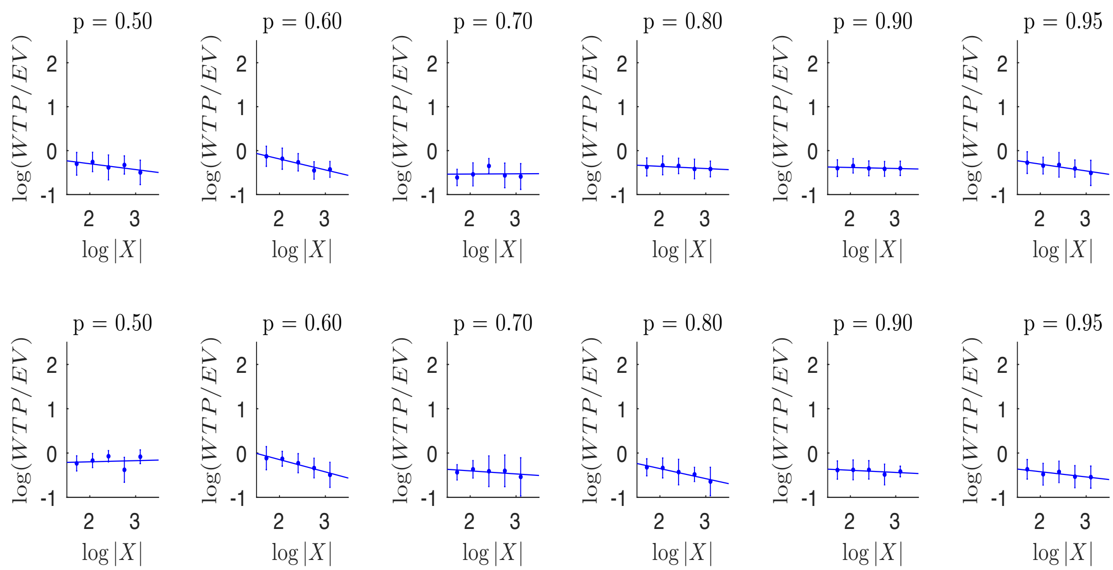

Figure 3: The same information as in Figure 2 (and using the same format), but now for probabilities $p \geq 0.50$.

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).[19] 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 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

[19] 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 $|X|$.

(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.[20]

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.[21] 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 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.[22] Here we again make the same assumption, and as in the

[20]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.

[21]The decision on some trials not to bid plainly does 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 2.6 below.

[22]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

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. (We make a specific hypothesis about the nature of this dependence, discussed below.)

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.[23] There is reason to believe that the discriminability between nearby numbers decreases in approximately this way as 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),[24] 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.[25] The same hypothesis about the noisy internal representation of numerical magnitudes is also consistent with observed biases in estimates of the numerosity of presented stimuli, as discussed in some detail in Khaw *et al.* (2021).

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,

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.

[23] Note that (2.1) implies that the probability that $r_{x2} > r_{x1}$ is an increasing function of $\log X_2 - \log X_1$. This is essentially the interpretation of Weber's Law (in other sensory domains) proposed by Fechner ([1860] 1966).

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

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

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 $[0, 1]$.[26]

There are a variety of reasons for choosing the specification (2.2) for the form of the encoding noise.[27] In the case of a lottery with two possible outcomes, what matters is the relative probability of the two outcomes occurring, not just the probability of one of them; and it would make sense for the relative odds to be represented in such a way that the precision of representation of the relative odds is the same in the case of two outcomes with probabilities $(1-p, p)$ as in the case of probabilities $(p, 1-p)$. In the case of small positive probabilities p of the larger outcome, the specification (2.2) is consistent with the idea that "Weber's Law" should hold for the discrimination of numerical magnitudes: it implies that people should more accurately distinguish a 2 percent probability from a 3 percent probability than they distinguish a 10 percent probability from an 11 percent probability (even though there is a difference of 1 percent in each case). But the specification (2.2) also implies, and to the same extent, that they should more accurately distinguish a 98 percent probability from a 97 percent probability than they distinguish a 90 percent probability from an 89 percent probability — which would not be implied if we were to assume simply that the quantity p were encoded logarithmically (as with the encoding of $|X|$).

Our specification is also consistent with the findings of Enke and Graeber (2022), 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,[28] 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.[29]

Studies of bias in the estimation of probabilities, relative frequencies, and proportions, when these are presented visually or through a sample of instances (rather than with number

[26]In addition, Zhang and Maloney (2012) argue that it is plausible to suppose that probability is represented in the brain in terms of log odds.

[27]Note that this model of noise in the encoding of probabilities was first proposed in Khaw *et al.* (2021).

[28]See the explanation in the Appendix, section A, 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 offered there, uncertainty about the certainty equivalent should be purely a reflection of the posterior uncertainty about p conditional on r_p. See also further discussion in section 5.1 below of the interpretation of elicited reports of subjective uncertainty.

[29]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 $p = 0.5$

symbols as in our experiment), support the view that (at least in these cases) the imprecision in people's recognition of probabilities or proportions have a similar property: their estimates are 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.[30] Our hypothesis here is that the pattern of imprecision in the internal representation of probabilities is the same when the probabilities are revealed to an experimental subject with number symbols, as when they must be discerned visually or by estimating the relative numbers of different elements in an array. This idea is parallel to our hypothesis above 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).

Indeed, our specific 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. 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 $\nu_z^2 = 2\nu_p^2$. Our conclusions below, however, depend only on assuming (2.2), and not on this particular interpretation of how the internal representation of the relative odds may be constructed.

2.3 Imprecise Response Selection

Our 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

[30] In the Appendix, section A, we show how the model (2.2) of noisy encoding of probability, together with a Bayesian model of how the noisy representation is "decoded" to produce an estimate, can give rise to a linear-in-log-odds pattern of estimation bias, of the kind shown by Zhang and Maloney (2012) to characterize many data sets.

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 $\mathbf{r} \equiv (r_p, r_x, \text{sign}(X))$, 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.[31] Because the 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.

Note that our model nests the case in which $\nu_c = 0$; in this case, our hypothesis would simply be that the bid C on any trial is the optimal one for the DM, 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). On the other hand, the hypothesis of random error in the selection of responses is common in models of perceptual judgments, and even more common in models of experimental data involving higher-level cognitive processing, like our lottery-valuation task. In the context of our theoretical model, the 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.[32]

Finally, 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

[31] 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.

[32] On the distinguishability of these different sources of error, in the case of both perceptual and cognitive judgments, see the review by Findling and Wyart (2021).

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 $\nu_c > 0$).

2.4 A Bayesian Decision Problem

We hypothesize that the function $f(\mathbf{r})$ is optimally adapted to the DM's situation. But a claim of optimality is necessarily relative to a particular objective, and to a particular class of possible decision situations. Here we further specify the DM's assumed objective, and the prior distribution over possible decision problems for which we assume that the DM's decision rule is optimized.

As stated in the introduction, our hypothesis is that the DM's decision rule is optimized to maximize the mathematical expectation of the DM's wealth, integrating over the set of possible decision situations that the DM might expect to encounter. Thus we must consider how a DM's stated willingness to pay should be expected to affect her financial wealth. 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 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

$$\begin{aligned} & W_0 \;+\; \frac{1}{N}\sum_i \mathrm{E}[I(B_i < C_i)\cdot(p_iX_i - B_i)] \\ = \; & W_0 \;+\; \frac{1}{N}\sum_i \mathrm{E}[I(B_i < p_iX_i)\cdot(p_iX_i - B_i)] \;-\; \frac{1}{N}\sum_i L(C_i;\, p_iX_i), \end{aligned} \tag{2.4}$$

where

$$L(C;\, V) \;\equiv\; -\int_V^C (V - B)\cdot g(B)dB. \tag{2.5}$$

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 (2.4) 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 (2.4). 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

$$\mathrm{E}[L(C;\, pX)\,|\mathbf{r}]. \tag{2.6}$$

If we suppose that the distribution $g(B)$ is approximately uniform[33] — that $g(B) \approx \tilde{g}$ over the relevant range,[34] where $\tilde{g} > 0$ is a constant — then we can approximate (2.5) as

$$L(C;\, pX) \;\approx\; \tilde{L}(C;\, pX) \;\equiv\; \tilde{g} \cdot \int_{pX}^{C} (pX - B)\, dB \;=\; \frac{\tilde{g}}{2}\,(C - pX)^2. \tag{2.7}$$

We use this approximation in our calculation of the numerical predictions of our model; that is, we assume that the DM's bidding rule minimizes the mean squared error that results from using C as an estimate of pX, the true expected value of the lottery.

If C could be chosen with precision, given an internal representation $\mathbf{r}$, the solution to this problem would be to choose

$$C \;=\; \mathrm{E}[pX\,|\mathbf{r}]. \tag{2.8}$$

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: (2.8) 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.

The posterior in the objective (2.6) depends on the prior distribution 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 X, the prior distribution for possible values of $|X|$ is of the form

$$\log |X| \;\sim\; N(\mu_x,\, \sigma_x^2), \tag{2.9}$$

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).

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

$$\log \frac{p}{1-p} \;\sim\; \text{Uniform}\,[\mu_z - \sqrt{3}\sigma_z,\; \mu_z + \sqrt{3}\sigma_z], \tag{2.10}$$

for some parameters μ_z, σ_z, which again indicate the mean and standard deviation of the prior.[35] Also, under the prior p and $|X|$ are distributed independently of one another (as is

[33]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 $g(B)$.

[34]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 (2.5).

[35]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.

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).[36]

2.5 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 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. 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 this amount that can be expressed in equivalent monetary units as

$$\kappa(\nu_x) \;=\; \tilde{A} \cdot \nu_x^{-2}, \tag{2.11}$$

where $\tilde{A} > 0$ is a parameter indexing the cost of greater precision in the representation of monetary amounts. The assumption of a cost that is linear in the precision 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 $\log |X|$[37] and an instantaneous variance $\sigma^2 > 0$ that is independent of $|X|$. 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.[38] 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 $|X|$.[39] 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 (2.11).

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

[36]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.

[37]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.

[38]Diffusion processes of this kind are often used to model the randomness in sensory perception and memory retrieval; see Gold and Heekeren (2014) for a review. Heng *et al.* (2022) 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.

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

complete representation **r** of the presented lottery, so as to minimize total expected losses

$$\mathrm{E}[\tilde{L}(C;\, pX) \,+\, \kappa(\nu_x(r_p))], \tag{2.12}$$

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.[40]

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 $(\widetilde{A}, \nu_z, \nu_c)$ that specify the degree of imprecision in internal representations. (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.

2.6 Declining to Bid

As already noted, on some 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.[41] 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.

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.[42] The prior distribution that is relevant for the "second-stage" problem modeled above (specified

[40]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 $\nu_x(r_p)$ in the text.

[41]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.

[42]See the Appendix, section D.3, for further discussion.

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.

3 Optimal Bidding and Stake-Size Effects

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 $A = \nu_z = \nu_c = 0$), 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.

3.1 Implications of the Logarithmic Model of Cognitive Noise

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 (2.9) 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[43]

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

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 the bidding.

This rule, together with (B.1), 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

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

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.[44]

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 (3.7) 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

[43]See the Appendix, section B.1, for details of the calculation.

[44]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.

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 Conformity of Our Data with the Model's Predictions

We now turn to quantitative tests of the degree to which our data conform to these predictions. 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

$$\log \frac{WTP}{EV} \sim N(m(p,X),\, v(p,X)). \tag{3.6}$$

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

$$m(p,X) = m(p,-X), \qquad v(p,X) = v(p,-X),$$

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" assumes that

$$m(p,X) = \alpha_p^+ + \beta_p^+ \log|X| \qquad \text{if } X > 0,$$
$$m(p,X) = \alpha_p^- + \beta_p^- \log|X| \qquad \text{if } X < 0.$$

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

$$m(p,X) = \alpha_p + \beta_p \log|X|, \tag{3.7}$$

regardless of the sign of $|X|$, for some coefficients (α_p, β_p) that depend only on the value of p. This model also imposes the restriction that $v(p,X)$ depends only on p and $|X|$, as it is a special case of the symmetric model. The "bounded symmetric affine model" imposes all of these restrictions, plus the further restriction that

$$0 \le \beta_p \le 1$$

for all p

Finally, we 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

Model	LL	BIC	K
	Pooled Data		
unrestricted model	-45579.8	92188.4	1
symmetric model	-45733.9	92067.9	1.47×10^{26}
general affine model	-45606.0	92008.7	1.05×10^{39}
symmetric affine model	-45746.1	91953.3	1.13×10^{51}
bounded symm. affine	-45746.1	91947.3	2.26×10^{52}
$\beta_p = 0$ for $p \geq 0.9$	-45751.8	91944.1	1.12×10^{53}
no stake effects	-45880.0	92145.3	2.29×10^{9}
	"Average Subject" Data		
unrestricted model	-1639.6	3602.7	1
symmetric model	-1654.6	3553.5	4.83×10^{10}
general affine model	-1642.2	3585.4	5.71×10^{3}
symmetric affine model	-1655.8	3521.6	4.08×10^{17}
bounded symm. affine	-1655.8	3518.8	1.65×10^{18}
$\beta_p = 0$ for $p \geq 0.3$	-1662.0	3505.0	1.64×10^{21}
no stake effects	-1671.0	3511.3	7.03×10^{19}
optimal bidding model	-1683.2	3397.6	3.44×10^{44}

Table 1: Measures of the goodness of fit of alternative statistical models of our subjects' responses. The first set of measures fit moments of the pooled data, while the second set fit moments of the "average subject's" responses. For each model, the log likelihood (LL) and Bayes Information Criterion (BIC) are reported, as well as the Bayes factor K by which each model is preferred to the unrestricted model.

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. Note that our model of optimal bidding implies that all of the restrictions of the bounded symmetric affine model should hold. We consider other models, however, in which some but not all of these restrictions are imposed, in order to study the particular ways in which our data do or do not conform to the predictions of our theoretical model. And we also consider more restrictive models in which β_p is required to equal zero for all large enough p, in order to provide quantitative measures of the importance of allowing for stake effects in order to match our data.

Table 1 reports measures of the goodness of fit of each of these models to the data on the responses of our subjects. The upper section of the table uses the pooled data from all 24 subjects to estimate the parameters of each of the model. Given that each of the models assumes a log-normal distribution of responses (3.6), 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 of the sample distribution of values for $\log WTP$. The likelihood also depends on N_j, the number of trials on which lottery j is evaluated. (See the Appendix, section C, 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.[45] 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.[46] 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).[47]

While the log likelihood is lower for more restrictive versions of the model, the BIC is lower — the greater parsimony more than outweighs the closeness of fit to the individual data moments — and as a result the more restrictive models have Bayes factors much greater than 1. In particular, the bounded symmetric affine model, imposing all of the general restrictions implied by our model of logarithmic coding, has a larger Bayes factor than any of the less-restrictive models; thus the data are more consistent with a characterization of this form.

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 for all large enough values of p; the largest Bayes factor is obtained if we set $\beta_p = 0$ for all $p \geq 0.9$. These are not restrictions implied by our model, which implies that $\beta_p < 0$ for all p. However, the fact that our data do not indicate values of β_p much lower than zero for high values of p (so that the greater parsimony of a model in which the zero coefficient is imposed for these values of p results in a lower BIC) does not disconfirm our model; for the model is consistent with these coefficients being only slightly negative, as we show below. The more important observation is that the data are *not* consistent with an assumption that $\beta_p = 0$ for all values of p, so that there are no stake-size effects at all. The best-fitting atheoretical model is one in which β_p is negative at least for low values of p.

The lower section of Table 1 reports similar measures of model fit for the same set of atheoretical characterizations of the data, but when instead of using the pooled data, we fit the models to the data for the "average subject." This means that for each lottery j, the value of $\hat{m}_j$ used to compute the likelihood of the data is the mean across subjects of each subject's mean $\log WTP$ for that lottery; the value of $(\hat{v}^j)^{1/2}$ used is the mean across subjects of each subject's standard deviation; and the value of N_j used is 8 (the number of times that each subject values a given lottery). We reach similar conclusions about the

[45] 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 the optimal bidding model (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.

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

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

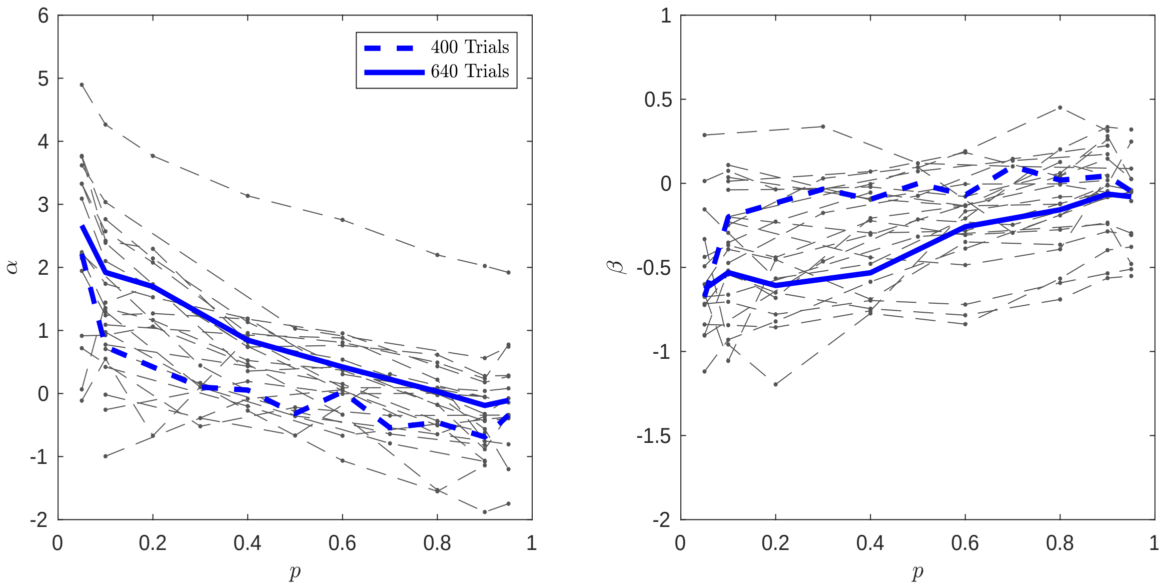

Figure 4: The coefficients $\{\alpha_p, \beta_p\}$ of the best-fitting symmetric affine model, estimated separately for each of our 24 subjects, and plotted as a function of p for each subject. The heavy curves indicate the median coefficients for each of two groups of subjects: the 12 who each completed 400 trials, and the 12 who each completed 640 trials.

relative goodness of fit of the different models using the "average subject" data, except that in this case the best-fitting atheoretical model is the bounded symmetric affine model in which $\beta_p = 0$ for all $p \geq 0.3$.[48]

Thus our experimental data support the general restrictions implied by our model, though we find that stake-size effects are notable only in the case of small values of p. The reason for the magnitude of the stake-size effects to depend on p is taken up in the next section.

3.3 Heterogeneity of Subject Responses

The atheoretical models just considered all assume either that a single set of coefficients $\{m_j, v_j\}$ should describe the valuations of all of our subjects, or that we are only interested in modeling the behavior of an "average subject." Yet there is also a fair amount of variation across subjects in the distribution of values elicited for a given lottery. For example, if we fit a symmetric affine model to the data for all subjects, but allowing the coefficients $\{\alpha_p, \beta_p\}$ and the residual variance v_j for each lottery to differ for each subject, the estimated coefficients for the different subjects vary considerably, as shown in Figure 4.

Some of the variation in these estimated coefficients may reflect over-fitting, given the small number of observations for each subject-lottery pair. However, even if we pool the 12 subjects who evaluated 400 lotteries in their session in one group, and the 12 subjects who evaluated 640 lotteries in another group,[49] and only allow the coefficients to differ between

[48] The bottom line of the table also reports measures of the goodness of fit for a model that imposes the further restrictions implied by our complete model of optimal bidding, discussed further below.

[49] As explained in the Appendix, section D.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

the two groups, we find evidence of heterogeneity in the behavior of the two groups.

While there is clearly variation in lottery valuations across subjects, we note that the general patterns of behavior identified in the pooled data (and 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 the theoretical prediction that $-1 < \beta_p < 0$ 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 also note a consistent pattern in the difference between the responses of subjects in the two groups: for all values of p, α_p tends to be larger, and β_p more negative, in the case of the subjects who undertook more trials. As we discuss below, this difference can be explained by our theoretical model, if we suppose that the longer session resulted in noisier internal representations in the case of the latter subjects.

4 Quantitative Implications of Endogenous Precision

We now discuss the further implications of our model, in the case of our specific assumptions about noisy encoding of information about the relative probabilities of the two outcomes, and about the way in which the precision of magnitude encoding depends on the perceived probabilities.

4.1 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 (3.7) 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 loss measure L defined in (2.7). This allows us to compute the conditional expectation $\mathrm{E}[\tilde{L}\,|\mathbf{r}]$

concentration). The groups also differ in the values of p used in the lotteries that they evaluated, though both groups faced both small and large values of p.

Integrating this over possible realizations of r_x (for a given value of r_p), we obtain an expression of the form[50]

$$\mathrm{E}[\tilde{L}\,|r_p] \;=\; Z(r_p) \;-\; \Gamma\varphi(r_p)\cdot\exp(\gamma_x(r_p)\sigma_x^2), \tag{4.1}$$

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 (4.1) 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 (2.11), we instead want to choose $\nu_x^2(r_p)$ to minimize the total loss

$$\mathrm{E}[\tilde{L}\,|r_p] \;+\; \kappa(\nu_x(r_p)) \tag{4.2}$$

associated with the internal representation r_p. (The objective (2.12) 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 (4.2) 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 B.2, that under the assumption that $\sigma_x^2 \leq 2$,[51] the solution to this optimization problem can be simply characterized. If

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

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

$$\frac{A}{(1-\gamma_x)^2} \;=\; \varphi(r_p)\exp(\gamma_x\sigma_x^2). \tag{4.3}$$

Equation (4.3) 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 (2.11) 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

[50] See the Appendix, section B.2, for details of the derivation.

[51] 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.[52] 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

$$\alpha_p \;+\; \beta_p \log \mathrm{E}[|X|] \;=\; \mathrm{E}\big[\log \mathrm{E}[p\,|r_p] - \log p\;|p\big] \;-\; \frac{3}{2}\nu_c^2. \tag{4.4}$$

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 $|X|$), (4.4) implies that the intercept will be a constant, the same for all p.[53] 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.[54] It then follows that when $|X| = \mathrm{E}[|X|]$, the sign of the intercept (4.4) 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

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

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

[54] See the Appendix, section A, for further discussion of these predicted biases in probability estimation.

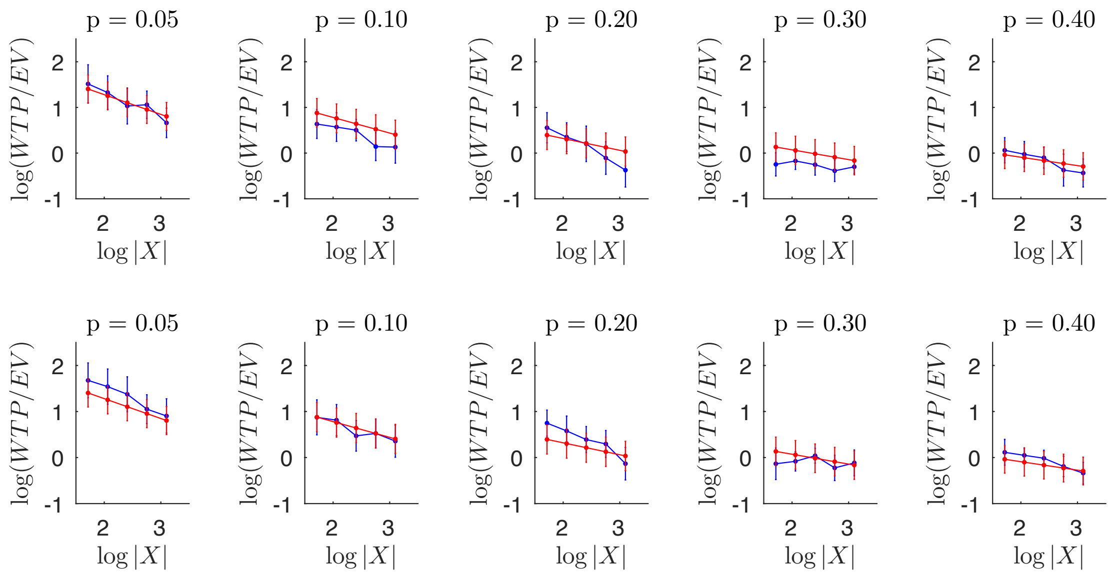

Figure 5: The same data as in Figure 2, but now compared with the predictions of the optimal bidding model with maximum-likelihood parameter estimates. (Blue: data for the "average subject." Red: theoretical predictions.)

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 consistent with both of the patterns documented in the previous literature.

4.2 Conformity of Our Data with the Model's Predictions

We test the conformity of our "average subject" data with the quantitative predictions of our model, 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 3.2, we write the likelihood as a function of the moments $\{\hat{m}_j, \hat{v}_j\}$ for the various lotteries j, and the number of trials N_j on which each lottery j 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 distribution.[55]

The theoretical moments $\{m_j, v_j\}$ 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,[56] but also on the parameters $(\mu_z, \sigma_z, \mu_x, \sigma_x)$ specifying the prior distribution over possible lotteries. Thus

[55]Recall that we assumed such a log-normal distribution (3.6) in the case of our atheoretical data characterizations. This is only an approximation in the case of our Bayesian model of optimal bidding on the basis of noisy internal representations. 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 not be exactly log-normal. Our log-normal approximation is discussed further in the Appendix, section C.

[56]Note that the composite parameter A, rather than the quantity $\tilde{A}$ appearing in (2.11), is the measure of the cost of precision in the encoding of numerical magnitudes that can be inferred from our behavioral data.

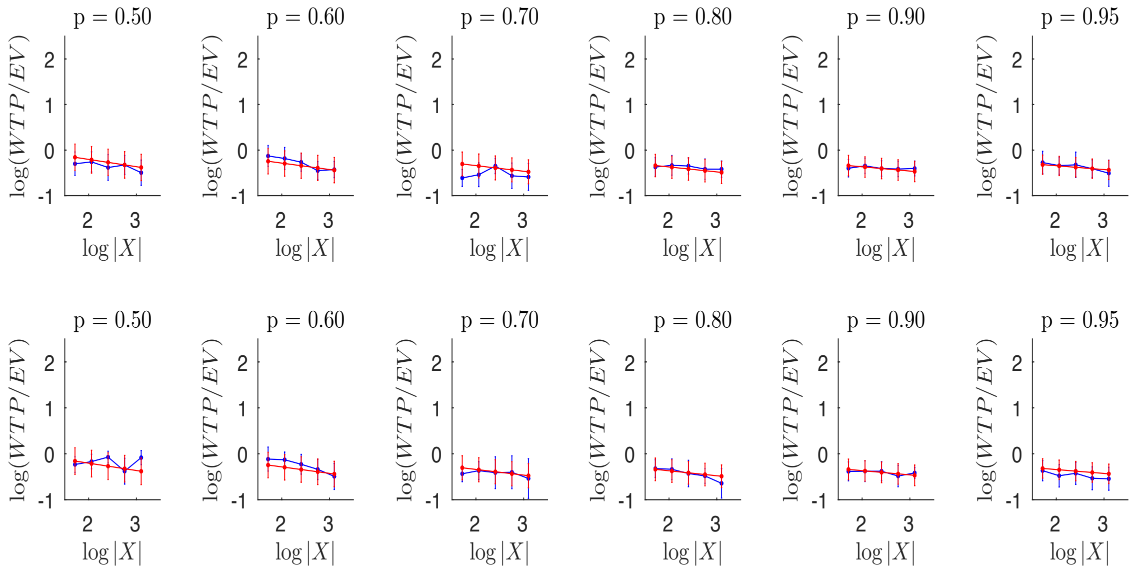

Figure 6: Continuation of Figure 5 for probabilities $p \geq 0.50$.

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 encoding 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.[57] 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 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^{23}).

Here we have fit the model parameters to the data moments for an "average subject," but as noted above, there is clearly 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 model parameters separately for two different "average subjects," one based on the 12 subjects who each evaluated 400 lotteries, and the other based on the 12 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.)

[57]The maximum-likelihood parameter estimates for the cognitive noise parameters are shown on the bottom line of the upper part of Table 2.

Alternative Parameter Estimates					
data	A	ν_z^2	ν_c^2	LL	LL/N
400-trial avg. subject	0.001	1.21	0.17	-1273.8	-3.195
640-trial avg. subject	0.015	2.54	0.16	-1773.5	-2.830
both average subjects	0.006	2.07	0.16	-3072.3	-2.996
single average subject	0.006	1.98	0.14	-1683.2	-3.283

Alternative Models of Both Average Subjects			
model	LL	BIC	K
common parameters	-3072.3	6179.3	1
separate parameters	-3047.3	6156.8	76,900

Table 2: Alternative estimates of the cognitive noise parameters for the optimal bidding 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 2 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 model considered on the bottom line of Table 1). 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). However, despite the differences in the bidding behavior of the two groups, the optimal bidding model does fairly well at explaining the behavior of each group; the value of LL/N is higher for *both* of the individual average subjects than in the fit to the data of a single overall average subject shown in Figures 5 and 6. The model fits best (in the sense of achieving a high value of LL/N) in the case of the 640-trial average subject. This may reflect a greater degree of heterogeneity within the 400-trial group, which could be due to the fact that they do not all face the same distribution of values of p.[58]

The bottom part of Table 2 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 (here treated as the baseline) is over 75,000.

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 a response to cognitive noise.

5 Extensions of the Basic Model

Here we briefly discuss how the theoretical model set out in section 2, and analyzed above, can be extended in straightforward ways that would allow our theory to be compared with other types of data.

5.1 Subjective Uncertainty and Risk Attitudes

Like us, Enke and Graeber (2022) propose that the fourfold pattern of risk attitudes of Tversky and Kahneman (1992) can be explained as an efficient response to cognitive noise

[58]See the Appendix, section D.1, for details. It might be appropriate to model different members of the 400-trial group as optimizing their decision rules for different priors; but we leave this for future investigation.

that leaves the DM uncertain about the correct valuation of a given lottery. A striking piece of evidence that Enke and Graeber offer in support of this view is a demonstration that experimentally elicited estimates of subjects' degree of uncertainty about how much to bid correlate with the degree to which their bids deviate from risk-neutral bids — in the direction predicted by prospect theory, and in each of the four quadrants of Tversky and Kahneman's "fourfold pattern." (That is, greater reported subjective uncertainty is associated with a larger *positive* relative risk premium in the gain domain, for all large enough values of p; with a larger *negative* relative risk premium in the gain domain, for all small enough values of p; and so on.) Even more strikingly, they show that the relationship is causal: an exogenous manipulation that should reduce the precision of subjects' awareness of the value of p is shown *both* to increase reported uncertainty about the value of the lottery and to result in larger deviations from risk-neutral valuations, again in all four quadrants.

Prospect theory predicts deviations from risk-neutral valuations with these respective signs, but it doesn't explain why the size of the deviations should be connected to uncertainty, or should be changed by Enke and Graeber's exogenous manipulation of uncertainty. Instead, our theory implies that observed deviations from risk-neutral valuations should all be attributed to the way in which it is optimal for subjects to shade their bids owing to the existence of cognitive noise, so that the predicted deviations should be larger when the cognitive noise is greater. If we suppose that subjects' reported subjective uncertainty about their bids is a monotonic function of the degree to which their bids will vary randomly from trial to trial on repeated considerations of the same lottery, then our model implies that an exogenous increase in the imprecision of internal representation of the value of p (an increase in the parameter ν_z^2) should both increase reported uncertainty and increase the degree to which the posterior mean estimate $\mathrm{E}[p\,|r_p]$ differs on average from the true value of p — both when the bias is positive and when it is negative.[59] This means that exogenous variation in cognitive noise (especially, noise in the internal representation of probabilities), whether due to continuing differences in individuals' cognitive capacities or to differences across experimental treatments, should result in systematic co-variation between subjective uncertainty and the strength of measured risk attitudes of the kind that Enke and Graeber document.

However, this interpretation of the results of Enke and Graeber (2022) depends on assuming that subjects' reported degree of subjective uncertainty should be connected to how unpredictable their behavior is according to our model. Enke and Graeber present evidence that it is, show that their subjects exhibit greater trial-to-trial variation in their valuations in the cases where they express greater subjective uncertainty about the right bid on any given trial (though they examine the issue of trial-to-trial variation in less detail than we do here). But is such a relationship consistent with our model? The model expounded in section 2 actually says nothing about what subjects' expressed degree of subjective uncertainty should be. It assumes that on any single trial, the DM's decision is based on a single draw of the internal representation $\mathbf{r}$ from a probability distribution of possible representations; this internal representation is random, according to our model, but not necessarily *known* to be random by the DM.[60] Our model implies that the decision process will result in different

[59]See the Appendix, section A, for further analysis of the way in which our theory can explain the findings of Enke and Graeber (2022).

[60]It should be recalled that we do not assume that the DM actually consciously carries out the Bayesian

outcomes upon repeated consideration of the same lottery valuation problem, but that need not be evident to the DM on any one of those occasions, since we do not assume any ability to recognize that the same lottery has already been assigned a value previously.

The results of Enke and Graeber indicate that subjects are capable of estimating their degree of uncertainty in a way that is at least somewhat correlated with the actual degree of variability of their cognitive process; in addition, it seems that when a change in the experimental design increases the degree of their cognitive imprecision, they are able both to recognize this (reporting a greater degree of subjective uncertainty) and to adapt their decision rule appropriately (changing their measured risk attitude). This requires that we assume a cognitive process with access to more information than is contained in the representation (r_p, r_x) discussed in section 2. One possibility is that in addition to an internal state parameterized by the two numbers (r_p, r_x), the DM's thought processes have access to other cues about the nature of the decision problem. For example, when the value of p is revealed in a more complex way (that requires an arithmetic calculation to determine the value of p), rather than by simply showing a number on the screen, the DM is aware of this difference — and thus, in addition to being aware of the value of the summary statistic r_p, they are also aware that the parameter ν_z^2 should be larger. We might then suppose that they are able to learn how to bid taking into account the cues about the current value of σ_z^2 as well as the elements of $\mathbf{r}$, and also able to learn how to answer questions about their subjective degree of uncertainty using this information.

Alternatively, we might suppose that the elements of the internal representation $\mathbf{r}$ are themselves higher-dimensional. We have justified the cost function (2.11) by supposing that $|X|$ is actually encoded by a series of independent noisy readings of $\log |X|$, with a cost proportional to the number of such readings. Our analysis in section 2 assumes that only the cumulative sum (or average) of these noisy readings is used as an input to the decision process, so that the information about $|X|$ can be summarized by a single real number, as assumed in (2.11). But if we assume that the decision process actually has access to the entire sequence of independent noisy readings, then it should be possible to extract an estimate of the variability of these readings from one to another, in addition to their average, and so to have access to an estimate of ν_x^2. One might assume the same about the internal representation of probability information; the normal distribution posited in (2.2) might represent the distribution of the average of a long sequence of independent noisy readings of the log odds, and thus one might assume that the decision process should have access to an estimate of ν_z^2 in addition to the value of the summary statistic r_p.

The interpretation of the results of Enke and Graeber (2022) that we propose here is similar to their own interpretation of the connection between cognitive uncertainty and valuation biases. One notable difference, however, is that Enke and Graeber interpret their elicited estimates of subjects' uncertainty about lottery values as reflecting metacognitive awareness of noise in the process by which they combine information about p and X to produce an estimate of EV (for example, noise in a process of mental multiplication), that need not be connected to any uncertainty about the values of p and X themselves.

calculations that we explain above; the DM need not be aware of entertaining a particular prior, and having particular posterior beliefs after conditioning on a noisy internal representation of the current decision problem, and so on.

We have instead modeled the sources of cognitive noise as noise in the retrieved internal representations of p and $|X|$,[61] that are then inputs to the DM's valuation rule, rather than any noise in the valuation rule itself (apart from the response noise parameterized by σ_c^2).

We have shown that we can explain the results of Enke and Graeber (at least qualitatively) on the basis of only the kinds of cognitive noise that we consider. We have also shown, in the case of our own experiment, that we can identify the separate signatures of the different types of cognitive noise that we allow for, and estimate the quantitative significance of each one in an estimated model of our subjects' behavior. But this does not answer the question whether computation noise of the kind proposed by Enke and Graeber might not also be important in explaining the behavior of our subjects. Our model is more parsimonious because of our neglect of this possibility, but we believe that the question of the extent to which alternative possible types of cognitive noise can be empirically distinguished deserves further study.

5.2 Choices Between Lotteries

We have presented a theory of how a Bayesian decision maker subject to cognitive noise should choose the amount that they are willing to pay for a lottery (p, X), when allowed to freely choose a bid, as in our experiment. Another common experimental design, however, involves the experimenter presenting the DM with a specific amount of money C that they can have with certainty, as an alternative to the risky lottery (p, X).

This method has been used in a number of the classic demonstrations of stake-size effects. One can compare decision problems in which both p and the ratio X/C remain fixed across problems, but the absolute size of C and X change, and ask whether the probability of choosing the risky lottery is affected by the change in stake size. Both Hershey and Schoemaker (1980) and Scholten and Read (2014) use this method, and find that increasing stake size reduces the probability of choosing a risky gain, but increases the probability of choosing a risky loss. We wish to consider whether our theory can also explain results of this kind.

Let us suppose that the features of the risky lottery (p, X) are encoded with noise in the way specified above; and let us correspondingly suppose that the magnitude of the certain amount $|C|$ is internally represented by a quantity r_c drawn from a distribution of the form

$$r_c \;\sim\; N(\log|C|,\, \nu_c^2), \tag{5.1}$$

by analogy with (2.1). As above, we simplify our analysis by assuming that the quantities r_p, r_x, r_c are all distributed independently of one another, conditional on the true data (p, X, C). And we suppose that the DM's decision can be based on perfect awareness that the probability of receiving the amount C will be 1, if that option is chosen, just as we assume

[61] Note that we interpret the random representations (r_p, r_x) as noisy *retrieved* values that are accessed as inputs to the valuation process. Our model does not require that subjects fail to correctly perceive the numbers that are presented to them on the screen, or even that they must be unable to subsequently answer questions correctly about the numbers that had been presented. It simply requires that the decision process not have access to any more precise internal representations of these quantities that might have existed at an earlier stage of mental processing, or in a part of their brain that is used for symbol-processing.

that there are no mistakes about whether gains or losses are at stake.[62] Since there is no variation across trials in the probability associated with the outcome C, or in its internal representation, the encoding noise parameter ν_c^2 is assumed to be the same on all trials.

Let us further suppose that the DM must choose between the risky lottery and the certain amount on the basis of these internal representations (r_p, r_x, r_c) and the common sign of X and C. The optimal decision rule will depend on the prior distribution over possible decision problems for which it is optimized. We assume a log-normal prior

$$\log |C| \;\sim\; N(\mu_c,\, \sigma_c^2),$$

by analogy with (2.9), and as above we simplify by assuming that p, X and C are distributed independently of one another. The optimal decision rule (that is, one that maximizes the expected value of the DM's financial wealth) will then be one that chooses the risky lottery if and only if

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

as in Khaw *et al.* (2021).

Let us consider the case in which X and C are positive (though our results can be directly extended to the case of losses as well). Then the methods used above allow us to express the optimal decision rule in the form

$$\log \mathrm{E}[p\,|r_p] \;+\; (1-\gamma_x(r_p))\bar{\mu}_x \;+\; \gamma_x(r_p)\cdot r_x \;>\; (1-\gamma_c)\bar{\mu}_c \;+\; \gamma_c \cdot r_c, \tag{5.2}$$

where we define

$$\gamma_c \;\equiv\; \frac{\sigma_c^2}{\sigma_c^2+\nu_c^2}, \qquad \bar{\mu}_c \;\equiv\; \mu_c \;+\; \frac{1}{2}\sigma_c^2,$$

by analogy with the definitions of γ_x and $\bar{\mu}_x$. The probability of choosing the risky lottery, in the case of a given problem (p, X, C), can then be derived as the probability that inequality (5.2) will hold when the random variables (r_p, r_x, r_c) are drawn from the conditional distributions specified in (2.2), (2.1), and (5.1).

As a special case, suppose that the same value of p is used on all trials (as in the experiment of Khaw *et al.*, 2021), and that the decision rule is optimally adapted to this prior. In such a case, the posterior mean $\mathrm{E}[p\,|r_p]$ is always equal simply to p, the value used on all trials; and there will be a single value for ν_x^2 (and hence for γ_x^2) on all trials. In this case, the probability that (5.2) will be satisfied can be computed from the properties of normal distributions. The predicted probability of choosing the risky lottery in a given problem is then

$$\mathrm{Prob}[risky] \;=\; \Phi\Big(\frac{\gamma_x \log X - \gamma_c \log C - q}{[\gamma_x^2\nu_x^2 + \gamma_c^2\nu_c^2]^{1/2}}\Big), \tag{5.3}$$

where

$$q \;\equiv\; (1-\gamma_c)\bar{\mu}_c \;-\; (1-\gamma_x)\bar{\mu}_x \;-\; \log p$$

[62]This assumption is consistent with the model of imprecise internal representation of probability information (2.2) proposed above. In that model, the case of a certain outcome lies at an infinitely distant point on the log-odds scale, so that adding a finite-variance Gaussian error term to the true log odds should yield an internal representation that does not have a positive likelihood under any probability $p < 1$. Bayesian inference from such an internal representation would then be equivalent to knowing that $p = 1$.

and $\Phi(z)$ is the CDF of a standard normal random variable.

If we further assume, as in Khaw *et al.* (2021), that the degree of imprecision with which both monetary payoffs X and C are encoded is the same ($\nu_c^2 = \nu_x^2$), and also that the degree of prior uncertainty about both quantities is the same ($\sigma_c^2 = \sigma_x^2$), then $\gamma_c = \gamma_x$, and (5.3) reduces to

$$\text{Prob}[risky] \;=\; \Phi\Big(\frac{\log(X/C) - (q/\gamma)}{\sqrt{2}\nu}\Big).$$

In this case, the model implies that there should be no stake-size effects: choice frequencies should depend only on the ratio X/C, as indeed is true to a good degree of approximation in the binary choice data of Khaw *et al.* (2021).

However, it need not always be the case that $\gamma_c = \gamma_x$. This assumption makes sense in the setting of Khaw *et al.* (2021), since if it is expected that the range of values offered as certain payoffs C is similar to the range of variation in the expected values pX of the risky lotteries, then one should expect the variance of $\log C$ to be similar to the variance of $\log(pX)$, which is just the variance of $\log X$ when (as in their experiment) p is always the same. But in an experiment in which p varies widely across trials, the same reasoning should lead one to assume that the variance of $\log C$ should be substantially greater than the variance of $\log X$, and hence that $\sigma_c^2 > \sigma_x^2$.

Then if we continue to assume (purely to simplify discussion) that ν_x^2 should be independent of r_p, and that $\nu_c^2 = \nu_x^2$, we should expect $\gamma_x < \gamma_c < 1$. It follows that if we increase stake sizes for any fixed ratio X/C, then for any internal representation r_p, the probability of satisfying (5.2) will be smaller the larger are the stakes. Then integrating over the distribution of representations r_p that may result from a given true probability p, we find that the probability of choosing the risky lottery will be smaller the larger are the stakes, for given values of p and X/C. (Using the same reasoning, but assuming that X and C are both negative, we find that the probability of choosing the risky lottery should be larger the larger are the stakes.) Thus the model can (at least qualitatively) explain the existence of stake-size effects of the kind reported by Hershey and Schoemaker (1980) and Scholten and Read (2014).

If we consider the limiting case in which σ_c is made unboundedly large (corresponding, essentially, to a prior under which any value of $\log C$ is considered equally likely), then $\gamma_c \to 1$, $(1-\gamma_c)\sigma_c^2 \to (1/2)\nu_c^2$, and (5.2) reduces to

$$\log \mathrm{E}[p\,|r_p] \;+\; (1-\gamma_x(r_p))\bar{\mu}_x \;+\; \gamma_x(r_p)\cdot r_x \;>\; r_c \;+\; \frac{1}{2}\nu_c^2,$$

or equivalently to

$$r_c \;<\; f(r_p, r_x) \;+\; \nu_c^2,$$

where $f(r_p, r_x)$ is the optimal bidding rule (3.3).

Thus conditional on any noisy internal representation (r_p, r_x) of the risky lottery, there is a close connection between the value of C required for the DM to be equally likely to choose the risky lottery or to decline it, in the binary-choice problem, and the value of C that will be the DM's median bid in the lottery-valuation problem.[63] It follows that the stake-size effects observed in our analysis of the lottery-valuation problem are precisely the ones that should also be observed in binary choices between a risky lottery and a particular

[63]For all values of (r_p, r_x), these two values of C differ by a constant multiplicative factor, $\exp(\nu_c^2)$.

certain amount, in the particular case of extreme agnosticism about the likely value of the certain amount that will be offered as an alternative. The scale-invariance of choice behavior observed in the experiment of Khaw *et al.* (2021) should be expected only in an environment that makes a different special kind of prior regarding the possible values of C likely, namely, the case in which prior uncertainty about $\log C$ on each trial should be essentially the same as prior uncertainty about $\log X$.

6 Conclusion

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, 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.

The prediction of stake-size effects is not the only respect in which our theory extends the predictions of prospect theory. Models in the spirit of prospect theory provide no reason for either the degree of stake-sensitivity or risk attitudes more generally to co-vary with the degree of variability of subjects' responses. Our model, which implies that departures from risk-neutral valuations should occur only as an adaptation to the presence of cognitive noise, instead implies that they should be closely connected. The results of our experiment provide support for this view, since the cases in which one observes the strongest departures from risk-neutrality (both the kind predicted by prospect theory and stake-size effects) are also the ones in which trial-to-trial responses are noisiest, namely the low-p lotteries (in both the gain and loss domains). We believe that future work should pay greater attention to the way in which apparent preferences co-vary with measures of cognitive imprecision.[64]

[64] In addition to the variability of responses across trials in the case of a given decision problem, that we emphasize in this study, such measures could include the subjective estimates of uncertainty elicited by Enke and Graeber (2022), or response times as in Alós-Ferrer *et al.* (2021).

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

[3] Bhui, Rahul, and Yang Xiang, "A Rational Account of the Repulsion Effect," *PsyArXiv* preprint, posted January 1, 2022.

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

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

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

[7] Enke, Benjamin, and Thomas Graeber, "Cognitive Uncertainty in Intertemporal Choice," NBER Working Paper no. 19577, December 2021.

[8] Enke, Benjamin, and Thomas Graeber, "Cognitive Uncertainty," working paper, Harvard University, May 2022.

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

[10] 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).

[11] 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).

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

[13] Frydman, Cary, and Salvatore Nunnari, "Cognitive Imprecision and Strategic Behavior," working paper, U.S.C., August 2022.

[14] Gabaix, Xavier, and David Laibson, "Myopia and Discounting," NBER Working Paper no. 23254, March 2017.

[15] Garcia, Miguel Barretto, Gilles de Hollander, Marcus Grueschow, Rafael Polania, Michael Woodford, and Christian C. Ruff, "Individual Risk Attitudes Arise from Noise in Neurocognitive Magnitude Representations," preprint, posted on *bioRXiv*, August 22, 2022.

[16] 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.

[17] Heng, Joseph A., Michael Woodford, and Rafael Polania, "Temporal Efficiency in Human Numerosity Estimation," working paper, ETH Zurich, May 2022.

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

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

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

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

[22] 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).

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

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

[25] 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).

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

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

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

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

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

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

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

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

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

[35] Thaler, Michael, "Overinference from Weak Signals, Underinference from Strong Signals, and the Psychophysics of Interpreting Information," working paper, Princeton University, September 2021.

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

[37] 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).

[38] 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).

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

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

[41] 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).

[42] 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).

ONLINE APPENDIX

Khaw, Li, and Woodford, "Endogenous Cognitive Imprecision and the Fourfold Pattern of Risk Attitudes"

A A Bayesian Model of the Estimation of Probabilities or Relative Frequencies

We have remarked in the main text that one justification for our interest in the model (2.2) for the noisy internal representation of the relative probabilities of the two outcomes derives from the consistency of such a model with the facts summarized in Zhang and Maloney (2012) regarding characteristic biases in people's estimates of probabilities or relative frequencies.

Zhang and Maloney 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

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

for some "anchor" or reference probability p_0 and an adjustment coefficient that in most cases satisfies $0 < \gamma < 1$. This implies a "conservative" bias in the estimates: probabilities are over-estimated when they are smaller than the reference probability and under-estimated when they are larger; thus in either case they are estimated to be closer to the reference probability than is actually the case. 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 that the model of Bayesian inference from a noisy internal representation of relative probabilities of the form (2.2) can explain not only the existence of a conservative bias, but the existence of a relationship (A.1) that is linear in the log odds, and variation in the reference probability p_0 depending on the range of probabilities p used in the experiment.[65] Suppose, as proposed in the main text, that the relative probability (or relative frequency) of the two outcomes is represented internally in a way that can be summarized by a real number r_p, that on any given trial will be a draw from a probability distribution of the form

$$r_p \;\sim\; N\Big(\log \frac{p}{1-p}, \, \nu_z^2\Big),$$

where p is the true probability or frequency, as indicated by the evidence presented to the subject.

[65] 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.

Bayesian decoding of this 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. 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,

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

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

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

where

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

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.[66] In this case, the model predicts an estimate

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

where the estimated log odds are given by $\hat{z}(r_p) = \hat{\mu}_z(r_p)$, the function defined in (A.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 (A.5).

Then since $\hat{\mu}_z(r_p)$ is a monotonic function, and the estimate for p specified in (A.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 (A.1), in which

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

[66]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.

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 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

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

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

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

which is again a relation of the form (A.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 (A.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 (2022), 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 B.1), 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

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

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 (2022) find.

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 (A.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 (A.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.

B Noisy Coding and Lottery Valuation: Derivations

Here we explain the details of the derivation of the theoretical model sketched in the main text.

B.1 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

$$\log C \;=\; f(r_p, r_x) \;+\; \epsilon_c, \tag{B.1}$$

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

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

where we have used (B.1) 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 concave 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

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

Substitution of this solution into (B.1) 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

$$\mathrm{E}[\log(C/pX)\,|p,X] \;=\; \alpha_p \;+\; \beta_p \log X, \tag{B.3}$$

where the coefficients

$$\begin{aligned}\alpha_p \;&\equiv\; \mathrm{E}\big[\log \mathrm{E}[p\,|r_p] - \log p\,|p\big] \;+\; (1-\gamma_p)[\mu_x + \frac{1}{2}\sigma_x^2] \;-\; \frac{3}{2}\nu_c^2, \\ \beta_p \;&\equiv\; -(1-\gamma_p), \\ \gamma_p \;&\equiv\; \mathrm{E}[\gamma_x(r_p)\,|p]\end{aligned}$$

all depend on the value of p but are independent of X. (Note that, among other things, this solution implies equation (4.4) 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 (B.3)) 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 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 (B.1), 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, (B.3) 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.

B.2 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 (B.2) over the possible realizations of r_x to obtain

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

Thus we can write

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

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 $\nu^2(r_p)$. 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 is given by

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

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

$$\min_{\gamma_x} F(\gamma_x;\, r_p) \;\equiv\; \frac{\tilde{A}}{\sigma_x^2}\Big(\frac{\gamma_x}{1-\gamma_x}\Big) \;-\; \Gamma\varphi(r_p)\cdot\exp(\gamma_x\sigma_x^2). \tag{B.4}$$

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

$$\frac{A}{(1-\gamma_x)^2} \;>\; \varphi(r_p)\exp(\gamma_x\sigma_x^2), \tag{B.5}$$

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 (B.5), we see that

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

where we define

$$G(\gamma_x;\, r_p) \;\equiv\; \log A \;-\; \log\varphi(r_p) \;-\; 2\log(1-\gamma_x) \;-\; \gamma_x\sigma_x^2. \tag{B.6}$$

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

$$A \;<\; \varphi(r_p), \tag{B.7}$$

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 (B.7) is a sufficient condition for the existence of an interior solution to the problem (B.4) at some $0 < \gamma_x < 1$. Moreover, the function defined in (B.6) 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 (B.7) holds, there must be exactly one solution to the first-order condition

$$G(\gamma_x;\, r_p) \;=\; 0, \tag{B.8}$$

an equivalent way of writing condition (4.3) stated in the main text. (Condition (4.3) in the main text is just the requirement that (B.5) 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 (B.4).

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 (B.7), 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 (B.6) 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 (B.8) 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 (B.4) 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 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

$$\text{LL} \;=\; \sum_i \left[L_1(p_i, X_i) \;+\; L_2(y_i\,|p_i, X_i)\right], \tag{C.1}$$

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 B.

We can write (C.1) in the form

$$\text{LL} \;=\; \sum_j N_j L_j, \tag{C.2}$$

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 (C.1), writing

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

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 (2.9) and (2.10) for the prior imply that

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

for any p_j such that

$$\mu_z - \sqrt{3}\sigma_z \ \leq \log \frac{p_j}{1 - p_j} \ \leq \ \mu_z + \sqrt{3}\sigma_z. \tag{C.5}$$

(Here we have omitted certain additive terms in (C.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 (C.5), i.e., outside the support of the prior (2.10), 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

$$\mu_z - \sqrt{3}\sigma_z \ \leq \min_j \log \frac{p_j}{1 - p_j}, \qquad \mu_z + \sqrt{3}\sigma_z \ \geq \max_j \log \frac{p_j}{1 - p_j},$$

where the minimum and maximum are over the set of probabilities used in the experiment.[67] Subject to these constraints, we find values of the parameters that maximize the function LL, using expression (C.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

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

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

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

[67]As explained further in Appendix section D.1, these minimum and maximum probabilities are 0.05 and 0.95 respectively.

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

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

Note that in (C.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[68]

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

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 (C.2) – (C.4) and (C.7). This is the method used for the results in Table 1 using "pooled data." In these calculations, the common parameters $\{m_j,\, v_j\}$ are assumed to specify a model that applies equally to each of the 24 subjects in our study. In the second part of Table 1, we instead fit the parameters of each model to the data moments of a fictitious "average subject."

If we let $\{\hat{m}_j^h,\, \hat{v}_j^h\}$ be the sample means and variances of the bids of some subject h (any of the subjects who express valuations for lottery j), then the data moments used in the "pooled data" calculations can be written as

$$\hat{m}_j = \frac{1}{N_j}\sum_h N_j^h \hat{m}_j^h, \qquad \hat{v}_j = \frac{1}{N_j}\sum_h N_j^h(\hat{v}_j^h + (\hat{m}_j^h)^2),$$

where N_j^h is the number of non-zero bids on lottery j by subject h.[69] The data moments of the "average subject" are instead computed as

$$\hat{m}_j^{avg} = \frac{1}{N_j}\sum_h N_j^h \hat{m}_j^h, \qquad (\hat{v}_j^{avg})^{1/2} = \frac{1}{N_j}\sum_h N_j^h (\hat{v}_j^h)^{1/2}. \tag{C.8}$$

These are the moments plotted in Figures 2 and 3.

In the calculations reported in the lower part of Table 1, we estimate the parameters of each atheoretical model so as to maximize the log likelihood of these "average subject" data. We again compute LL for any model parameters $\{m_j,\, v_j\}$ using equations (C.2) – (C.4) and (C.7), but substituting $\{\hat{m}_j^{avg},\, \hat{v}_j^{avg}\}$ for the data moments in (C.7), rather than the pooled-data moments $\{\hat{m}_j,\, \hat{v}_j\}$. We similarly substitute the quantities $\{N_j^{avg}\}$ for the quantities $\{N_j\}$ in (C.2). Here N_j^{avg} is the effective number of observations of bids on lottery j by the average subject, defined as

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

[68]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.

[69]This can be less than 8 for some subjects, if they bid zero on some trials. See further discussion in Appendix section D of the exclusion of those trials from our data set.

where H_j is the number of subjects bidding on lottery j whose data are averaged in order to define the moments of the average subject.[70]

In the same way, we estimate the parameters of our optimal bidding model so as to maximize the log likelihood of the "average subject" data.[71] 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 (C.6); while (B.1) 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 (C.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 $\{\hat{m}_j^{avg}, \hat{v}_j^{avg}\}$.[72]

Our MLE estimates of the parameters of the optimal bidding model (reported in Table 2) are obtained by maximizing the approximate likelihood function calculated in this way. The reported values of LL and BIC similarly use the maximized value of the approximate likelihood function. And finally, the value of LL$/N$ reported in Table 2 actually divides LL by $N^{avg} = N/H$, the average number of bids per subject, where H is the number of subjects whose data are averaged.

The same method is used in Table 2 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, we compute data moments for a fictitious subject as in (C.8), but now 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.[73] In (C.8), 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;[74] 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.[75]

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

[70]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 3 below.

[71]We don't fit the optimal bidding model to the pooled data, since we don't think that it makes sense to attribute the variability of the bids for a given lottery, as measured by $\hat{v}_j$, entirely to randomness of the trial-to-trial bidding of a single type of subject.

[72]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."

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

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

[75]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 D.3. For the 640-trial average subject, N^{avg} is actually only equal to 626.67. This is why in Table 2, the number given for LL$/N$ is not equal to the number given for LL divided by 640.

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

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). Finally, 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.[76]

D Experimental Data: Additional Details

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

D.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 3. 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 3) 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.

Out of the 28 subjects listed in Table 3, four subjects were excluded from the study on the ground that their responses suggested limited attention to (or misunderstanding of) the assigned task. These were subjects 9, 10, 11, and 19, discussed further in the next subsection.

[76]Because the fraction of zero bids is smaller in the case of the 400-trial subjects, as discussed below, this results in $N^{avg} = 398.75$, a number only slightly less than 400.

With these exclusions, we are left with data from 24 subjects, 12 in the 400-trial group and 12 in the 640-trial group.

D.2 Excluded Subjects

We excluded four subjects from use in our analysis, on the ground that their bids seemed insufficiently sensitive to the data defining each individual lottery for us to think that they were taking much care to think about the problem presented. To quantify this, we decomposed the variance of each subject's bids into two parts:

$$\mathrm{var}[\log WTP] \;=\; \sum_j \frac{N_j}{N} \mathrm{var}_j(\log WTP) \;+\; \mathrm{var}[E_j(\log WTP)],$$

where $E_j[\cdot]$ and $\mathrm{var}_j[\cdot]$ refer to mean and variance of the distribution of bids associated with a particular lottery j, and N_j/N is the fraction of all bids by that subject that are for lottery j. Thus the first term on the right-hand side measures the average variability of the subjects' bids for a particular lottery, while the second term measures the variability across lotteries of the subject's average bid. Our proposed measure of inattentiveness is then e, the fraction of the variance accounted for by the first term. (Note that we are not assessing the correctness of anyone's bids, or their conformity to any theory, but simply the degree to which the trial-to-trial variation in their bids is accounted for by the different lotteries that are presented on different trials.

For each subject, we separately compute the measure e for gain trials and loss trials. We then retain the subjects in our data set if and only if $e < 2/3$ for both gain and loss trials. All except subjects 9, 10, 11, and 19 pass this test.[77] We measure e separately for the gain and loss trials, because some subjects appear to be much more inattentive on one kind of trials.

Figure 7 gives information about the bidding of each of the four excluded subjects. Subject 11 is a particularly clear example of inattentiveness: the subject's distribution of slider positions is roughly the same for all lotteries, with little evident sensitivity to the values of either p or X on the given trial.[78] This behavior is not inconsistent with our model, of course, but the model would attribute a very large degree of cognitive noise to this subject's internal representations of both p and X, different from those of most other subjects.

Subjects 9 and 19 are instead examples of less apparent attention (or less understanding of the task) in the case of lotteries involving losses. Both of these subjects differentiate their bids in a fairly reasonable way in the case of lotteries involving gains. But subject 19 bids similarly for all lotteries involving losses (regardless of the values of p and $|X|$), suggesting confusion about the question asked in the case of lotteries involving losses. Subject 9's bids in the loss domain are relatively insensitive to the value of $|X|$, and while they do depend

[77]Subject 9 is insufficiently sensitive to the lottery data in the case of the gain trials; subjects 10 and 19 are insufficiently sensitive in the case of the loss trials; and subject 11 is highly insensitive to the lottery data on both kinds of trials.

[78]The figure shows the subject choosing bids around \$10 in the case of lotteries involving gains, and bids around -\$10 in the case of lotteries involving losses; but this is because of the different way in which we interpret the subject's slider position in the two cases. There is not much evidence that the subject pays attention even to the sign of the payments involved on a given trial.

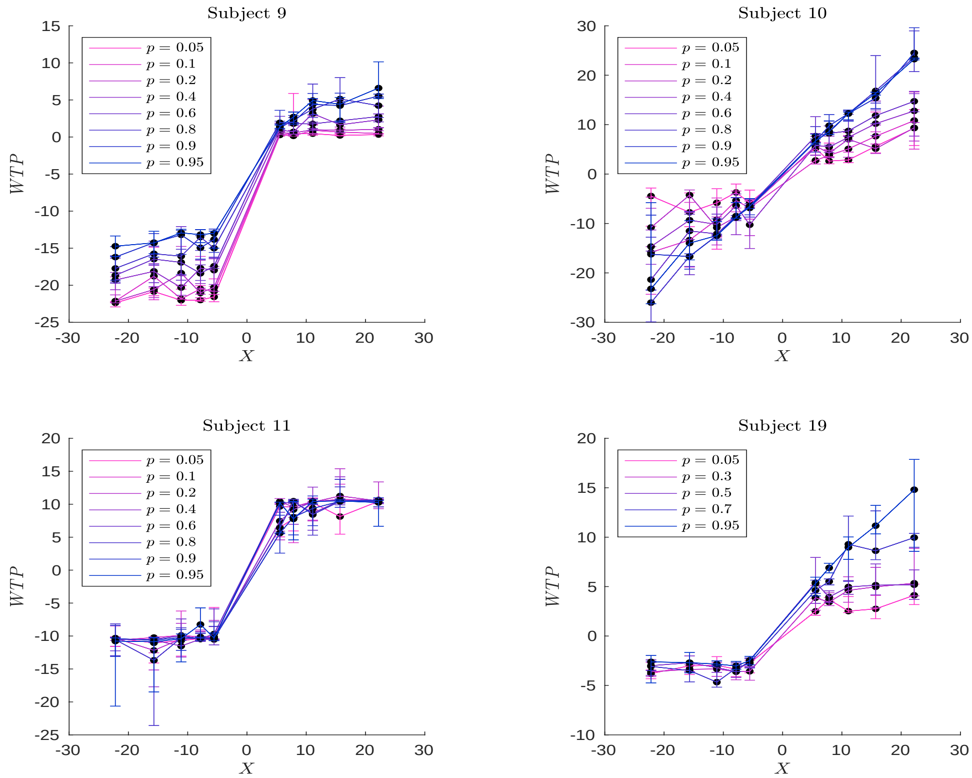

Figure 7: Bid distributions for each of the lotteries (as functions of X and p), for each of the four excluded subjects. Colors indicate the value of p. For each lottery, the dots (connected by lines for each p) indicate the median bid, and the upper and lower whiskers indicate the 25th and 75th-percentile bids for that subject.

on the value of p, they don't depend on p in a sensible way: the subject offers to pay *more* to avoid a random loss when the probability of the negative outcome is *smaller*. This again suggests confusion about the question being asked (perhaps confusion about the significance of different slider positions in this case).

Finally, subject 10's bids are somewhat sensitive (and in the reasonable direction) to the values of both p and $|X|$, in both the gain and loss domains; but subject 10's bidding behavior is noisier than that of the non-excluded subjects. If we fit the optimal bidding model to the data of subject 10 alone, the estimated cognitive noise parameters are much larger for subject 10 than for the average subject from among the non-excluded 640-trial subjects, as shown in Table 4.[79] Thus describing the behavior of subject 10 with separate parameters, rather than considering subject 10 in the group used to define the behavior of the 640-trial average subject, seems to be appropriate even if one might wish to also consider the fit of our model to the bidding by subject 10 when assessing the success of the model.[80]

[79] The parameter estimates and value for LL/N shown in the table for the average subject from among the 640-trial subjects (other than subjects 9, 10, and 11) are the same as the ones already reported in Table 2 in the main text.

[80] With regard to the fit of the optimal bidding model to the data of subject 10: when we allow for subject-specific parameters, the fit is no worse than the one shown in Figures 5 and 6. The value of LL/N is actually slightly larger when the model is fit to the data of subject 10, than in the case of the fit to the data of the

Alternative Parameter Estimates				
data	A	ν_z^2	ν_c^2	LL/N
640-trial (exc. 9, 10, 11)	0.015	2.54	0.16	-2.830
subject 10	1.03	6.73	0.21	-3.275

Table 4: Alternative estimates of the cognitive noise parameters for the optimal bidding model, for the "average subject" based on the data for 640-trial subjects other than subjects 9, 10, and 11, and when the model is fit to the data of subject 10.

Thus we offer the following summary of the differences between the data of the excluded subjects and those included in the data analyzed in the main text: Two of the excluded subjects (subjects 10 and 11) bid in a way that is consistent with our model of optimal bidding in the presence of cognitive noise, but would require larger values of the cognitive noise parameters to rationalize their behavior than in the case of the included subjects. The other two excluded subjects (subjects 9 and 19) do not behave in a way consistent with our model, but seem not to have understood the task (in particular, not to have understood the lotteries involving losses).

It is notable that the subjects who rate as least attentive according to our e measures are often the ones in the 640-trial group. Three out of four of the excluded subjects are from this group; but in addition, among the 15 subjects for whom $e > 1/4$ when all trials are considered, 12 are 640-trial subjects. Thus in general, the 640-trial subjects made bids that respond less precisely to the data presented to them, perhaps because of fatigue or boredom. This suggestion is also supported by the observation that the estimated cognitive noise parameters for the 640-trial average subject are larger than those for the 400-trial average subject.

D.3 Zero Bids

In addition to excluding the bids of four subjects altogether, we also drop the bids of the 24 remaining subjects when the bid is exactly 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 2 percent of all trials. Zero bids were instead relatively rare for the 400-trial subjects, who submitted only 15 such bids (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.[81] 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

"average subject" shown in Figures 5 and 6. (Compare the value on the bottom line of Table 4 with the one on the fourth line of Table 2.)

[81]Note that the data shown in the figure are only for the 24 non-excluded subjects.

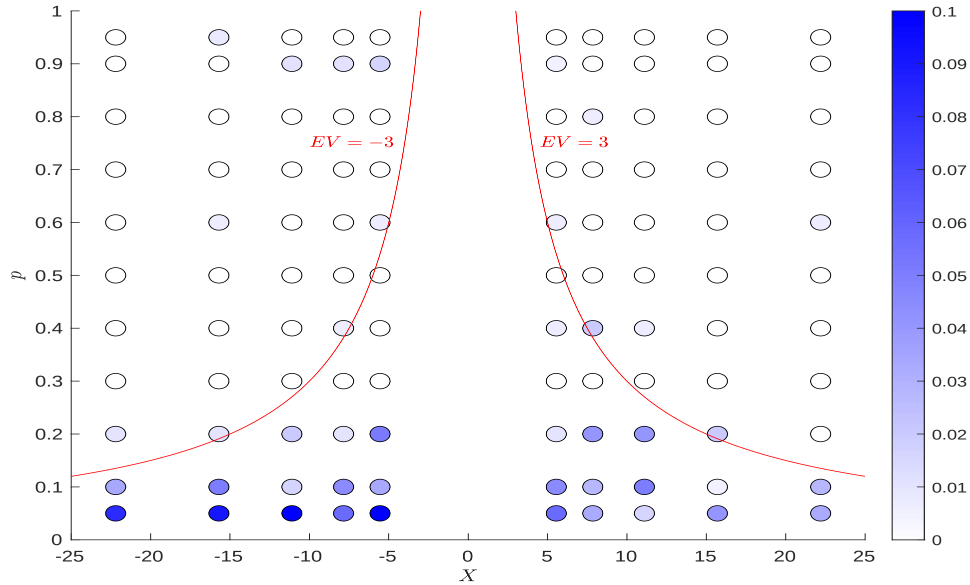

Figure 8: The fraction of zero bids for each of the lotteries (X, p) that are presented to subjects. (Color code is explained by the scale at the right.)

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.[82] Zero bids were especially common in the case of lotteries involving losses and only a small probability ($p = 0.05$) of a non-zero loss; in this case, zero bids were submitted on 8.7 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.[83]

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).[84] The DM's prior in the second stage (when it is reached) should then depend

[82] 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 provided by the responses of excluded subjects 9 and 19, shown in Figure 7.

[83] 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).

[84] An empirical measure of this probability is given by one minus the fraction shown in Figure 8 for each

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.

lottery (p, X).