Review of Economic Studies (2021) **88**, 1979–2013 doi:10.1093/restud/rdaa044

Advance access publication 3 August 2020

Cognitive Imprecision and Small-Stakes Risk Aversion

MEL WIN KHAW
Duke Institute for Brain Sciences, Duke University
ZIANG LI
Department of Economics, Princeton University
and
MICHAEL WOODFORD
Department of Economics, Columbia University

First version received May 2019*; Editorial decision April* 2020*; Accepted June* 2020 (*Eds.*)

Observed choices between risky lotteries are difficult to reconcile with expected utility maximization, both because subjects appear to be too risk averse with regard to small gambles for this to be explained by diminishing marginal utility of wealth, as stressed by Rabin (2000), and because subjects' responses involve a random element. We propose a unified explanation for both anomalies, similar to the explanation given for related phenomena in the case of perceptual judgments: they result from judgments based on imprecise (and noisy) mental representations of the decision situation. In this model, risk aversion results from a sort of perceptual bias—but one that represents an optimal decision rule, given the limitations of the mental representation of the situation. We propose a quantitative model of the noisy mental representation of simple lotteries, based on other evidence regarding numerical cognition, and test its ability to explain the choice frequencies that we observe in a laboratory experiment.

Key words: Weber's Law, random utility, prospect theory, Rabin critique

JEL Codes: C91, D03, D81, D87

Risk-averse choices are conventionally explained as reflecting expected utility maximization (EUM) on the part of decision makers for whom the marginal utility of additional wealth decreases with increases in their wealth. Experimentally observed choices under risk are difficult to reconcile with this theory, however, for several reasons.[1] One is the observation that people often decline even very small bets that offer somewhat better than fair odds. In the case of any smooth utility-of-wealth function, choices ought to become nearly risk-neutral in the case of small enough stakes (Arrow 1971). And while it is always possible to explain rejection of any given bet by assuming sufficient rapidly diminishing marginal utility of wealth, the degree of curvature of the utility function that is required will then imply that the same person should reject even extremely

1. For a broader review, see Friedman *et al.* (2014).

The editor in charge of this paper was Nicola Gennaioli.

favourable bets when potential losses are moderately large (though in no way catastrophic), as explained by Rabin (2000); this too seems plainly counter-factual.[2]

A well-known response to this difficulty (Rabin and Thaler 2001) is to propose that people maximize the expected value of a non-linear utility function, but that this function is *reference-dependent:* it is not a context-invariant function of wealth, but instead depends on how the wealth that may be obtained in different possible states compares to some reference level of wealth.[3] But this proposed solution raises a further question: why the human mind should exhibit such reference-dependence, if it leads to behaviour that would seem not to be in the decision maker's interest.[4] Simply stating that this appears to be what many people prefer—as if they perfectly understand what they are getting from their choices and nonetheless persistently choose that way—is not entirely convincing. We propose instead an alternative interpretation, under which decision makers often fail to accurately choose the option that would best serve their true objectives, because their decision is based not on the exact characteristics of the available options, but rather on an imprecise mental representation of them.

Our alternative explanation has the advantage that it can simultaneously explain another well-established feature of choice behaviour in experimental settings: that choices appear to be random, in the sense that the same subject will not always make the same choice when offered the same set of simple gambles on different occasions (Hey and Orme 1994; Hey 1995, 2001). We propose that this should be understood as reflecting imprecision in the cognitive processes involved in making a choice, in the same way as random trial-to-trial variation in perceptual judgments (say, about the relative magnitude of two sensory stimuli) is understood.[5]

As reviewed in Woodford (2020), the standard approach to modelling imprecision of this kind in perceptual judgments, since the work of Fechner (1860) and Thurstone (1927), attributes the randomness in perceptual judgments to randomness in internal representations of the stimulus features in the brain of the observing subject, and not necessarily to any sub-optimality of the judgments that are produced on the basis of that imperfect representation of the world. Such an approach has the advantage of allowing for random variation in the responses that are given on individual occasions, while nonetheless retaining a fairly constrained specification of the nature of the randomness. The "decoding" of the information contained in the noisy internal representation can be assumed to be optimal, leading to a precise specification of this part of the model on normative grounds; the nature of the noise in the "encoding" process is specified more flexibly, but this can be a subject of empirical study (*e.g.* using neurophysiological measurements), and can also be experimentally manipulated (by changing conditions in ways that should predictably increase or decrease encoding noise), as we discuss further below.

2. Rabin's argument appeals to introspection. But see Cox *et al.* (2013) for examples of experiments in which subjects make choices with respect to both small and large bets that are inconsistent with EUM under any possible concave utility function.

3. Fudenberg and Levine (2006, 2011) offer a different potential explanation. But as with the hypothesis of reference-dependent preferences, their explanation for small-stakes risk aversion provides no explanation for other phenomena that we address in this article, such as the observation of risk-seeking with respect to small losses along with risk-aversion with respect to small gains.

4. As Rabin and Thaler (2001) point out, "myopic loss-averters ... make decisions that mean that others can take money from them with very high probability and very low risk." They also note that such exploitation seems all too commonplace. Our point is not to assert that the predictions of such a model must be wrong, but rather to urge that persistent behaviour of this kind calls for an explanation.

5. For additional types of evidence also suggesting that the valuations that can be accessed when making a decision are imprecise, see Butler and Loomes (1988, 2007) and Bayrak and Hey (2019).

In essence, we propose that small-stakes risk aversion can be explained in the same way as perceptual biases that result from noise in internal representations.[6] According to our theory, intuitive estimates of the value of risky prospects (not ones resulting from explicit symbolic calculations) are based on mental representations of the magnitudes of the available monetary payoffs that are imprecise in roughly the same way that the representations of sensory magnitudes are imprecise, and in particular are similarly random, conditioning on the true payoffs. Intuitive valuations must be some function of these random mental representations. We explore the hypothesis that they are produced by a decision rule that is optimal, in the sense of maximizing the (objective) expected value of the decision maker's expected wealth, subject to the constraint that the decision must be based on the random mental representation of the situation.

Under a particular model of the noisy coding of monetary payoffs, we show that this hypothesis will imply apparently risk-averse choices: the expected net payoff of a bet will have to be strictly positive for indifference, in the sense that the subject accepts the bet exactly as often as she rejects it. Risk aversion of this sort is consistent with a decision rule that is actually optimal from standpoint of an objective (expected wealth maximization) that involves no "true risk aversion" at all; this bias is consistent with optimality in the same way that perceptual biases can be consistent with optimal inference from noisy sensory data. And not only can our theory explain apparent risk aversion without any appeal to diminishing marginal utility, but it can also explain why the "risk premium" required in order for a risky bet to be accepted over a certain payoff does not shrink to zero (in percentage terms) as the size of the bet is made small, contrary to the prediction of EUM.

This explanation has two important advantages over other proposed explanations. First, it is more parsimonious: rather than introducing separate free parameters to account for risk attitudes on the one hand and the randomness of choice on the other, the same parameter (the degree of noise in internal representations) must account for both phenomena in our theory. In addition, the same hypothesis of optimal choice on the basis of noisy internal representations provides a unified explanation for a number of additional experimentally observed phenomena (such as the "isolation effect" and "reflection effect" of Kahneman and Tversky 1979), that require additional, seemingly independent hypotheses in the account of them given in prospect theory.

And second, the hypothesis that these effects result from imprecision in internal representations, rather than from true preferences and random error in subjects' reporting of those preferences, suggests that the degree to which both stochastic choice and risk aversion are observed should vary across individuals and situations as a result of variation in the amount of working memory capacity that can be devoted to the representation of the numerical magnitudes involved in a given choice problem. Various pieces of evidence (discussed further in Section 5) suggest that this is the case.

Both our own results below and a replication of our work by Garcia *et al.* (2018) show that subjects whose choices are more random tend to exhibit greater small-stakes risk aversion, and Garcia *et al.* further show that both the randomness of choice and risk aversion in choices between lotteries correlates with an independent measure of imprecision in subjects' internal representations of numerical magnitudes. Moreover, subjects' capacity for representation of numerical magnitudes appears to be subject to experimental manipulation. Authors such as Gerhardt *et al.* (2016) find that increased "cognitive load" increases the apparent risk aversion of subjects, while Frydman and Jin (2019) show that choice between lotteries can be made less stochastic by reducing the range of payoffs used across different trials. None of these effects

6. Our theory is thus similar to proposals in other contexts (such as Koszëgi and Rabin 2008) to interpret experimentally observed behaviour in terms of mistakes on the part of decision makers—*i.e.*, a failure to make the choices that would maximize their true preferences—rather than a reflection of some more complex type of preferences.

can be explained by a preference-based account of either small-stakes risk aversion or stochastic choice.

While our approach is based on an analogy with imprecision in perceptual judgments, as modelled in the psychophysical tradition originating with the work of Fechner (1860), it differs from what are sometimes called "Fechnerian" models of stochastic choice (*e.g.* Block and Marschak 1960, or Butler *et al.* 2012). In such models, a deterministic valuation of a given lottery is assumed to be precisely computed (say, in accordance with expected utility theory), but then a random term is added to the valuation in any individual trial; such models allow choice to be stochastic, but risk attitudes must still be explained by the "deterministic core" of the model, in a way that is independent of the randomness. In our theory, instead, the randomness in choice is attributed to random errors at an earlier stage in the valuation process, which also give rise to departures from risk neutrality.

At the same time, suggesting an analogy with perceptual biases does not mean that the random errors of interest to us arise only in the perception of data as they are presented to the subject. This is not even true of perceptual judgments, except of the simplest kind. Drugowitsch *et al.* (2016) consider perceptual comparison tasks in which the properties of two sequences of stimuli must be averaged and then compared with one another, allowing them to determine (from the way in which the degree of randomness in responses varies with changes in the sequences to be compared) how much of the randomness in responses is due to (a) noise in the initial perceptions of individual stimuli, (b) noise in subsequent processing of the initial sensory data to obtain an assessment of the average for the sequence, or (c) noise in reporting the answer implied by those internal computations. They conclude that in their task, the most important source of randomness is of type (b), which the call "inference noise." Similarly in our case, we suppose that the important source of noise is error in the representations of numerical magnitudes that are subsequently accessed in the decision process, rather than error in the initial recognition of the meaning of the symbols presented to the experimental subject.

Under our hypothesis about the source of small-stakes risk aversion, it is important to jointly model the determinants of average responses and the randomness in the responses on individual trials. This has implications not only for theoretical modelling, but for the experimental procedures that should be used to measure risk attitudes. One should not expect to be able to measure people's noise-free preferences by simply measuring modal or median responses, and assuming that a deterministic model of choice should fit those summary data, as is often presumed in the experimental literature, as a way of side-stepping the issue of modelling the noise in individual responses. Instead, the perspective that we propose implies that it is necessary to model the entire distribution of responses that are predicted to occur under any given objective choice situation.

Given this, different experimental procedures are appropriate as well. Rather than seeking to obtain a single measure of what is "really" preferred in a given choice situation—by offering each choice problem only once, and presenting closely related problems together in order to encourage consistent answers across several problems—we instead use methods similar to those of perceptual studies (or papers like Mosteller and Nogee 1951): individual choice problems are presented repeatedly to the same subject, but in random order so as to encourage an independent judgment in each case. For this reason, we offer new experimental evidence, despite the extensive prior work on choice between simple gambles.

Section 1 reviews evidence regarding the mental representation of numerical magnitudes that motivates our model of noisy coding of monetary payoffs. Section 2 presents an explicit model of choice between a simple risky gamble and a certain monetary payoff, and derives predictions for the both the randomness of choice and the degree of apparent risk aversion implied by an optimal decision rule. Section 3 describes a simple experiment in which we are able to test some of the specific quantitative predictions of this model. Section 4 discusses the implications of our theory for additional phenomena reported in experiments such as those of Kahneman and Tversky (1979). Finally, Section 5 discusses further evidence suggesting that small-stakes risk aversion reflects imprecise cognition, rather than perfectly understood preferences, and concludes.

1. IMPRECISION IN NUMERICAL COGNITION

An important recent literature on the neuroscience of perception argues that biases in perceptual judgments can actually reflect optimal decisions—in the sense of minimizing average error, according to some well-defined criterion, in a particular class of situations that are possible ex ante—given the constraint that the brain can only produce judgments based on the noisy information provided to it by sensory receptors and earlier stages of processing in the nervous system, rather than on the basis of direct access to the true physical properties of external stimuli (*e.g.*, Stocker and Simoncelli 2006; Wei and Stocker 2015). The approach has been used to explain systematic biases in perception in a variety of sensory domains (Petzschner *et al.* 2015; Wei and Stocker 2017).

The relevance of these observations about perceptual judgments for economic decision might nonetheless be doubted. Some may suppose that the kind of imprecision in mental coding just discussed matters for the way in which we perceive our environment through our senses, but that an intellectual consideration of hypothetical choices is an entirely different kind of thinking. Moreover, it might seem that typical decisions about whether to accept gambles in a laboratory setting involve only numerical information about potential payoffs that is presented in an exact (symbolic) form, offering no obvious opportunity for imprecise perception. However, we have reason to believe that reasoning about numerical information often involves imprecise mental representations of a kind directly analogous to those involved in sensory perception.

1.1. *Imprecise perception of numerosity*

This is clearest (and has been studied most thoroughly) in the case of perceptions of the number of items present in a visual display. For example, quick judgments can be made about the number of dots present in a visual display of a random cloud of dots, without taking the time to actually count them. As with perceptions of physical magnitudes such as length or area, such judgments of numerosity are subject to random error. And just as in the case of sensory magnitudes, the randomness in judgments can be attributed to randomness in the neural coding of numerosity (Nieder and Merten 2007; Nieder and Dehaene 2009; Nieder 2013).

We can learn about how the degree of randomness of the mental representation of a number varies with its size from the frequency distribution of errors in estimation of numerosity. A common finding is that when subjects must estimate which of two numerosities is greater, or whether two arrays contain the same number of dots, the accuracy of their judgments does not depend simply on the absolute difference in the two numbers; instead, the absolute difference required for a given degree of accuracy grows as the numbers are made larger, and roughly in proportion to their magnitudes—a "Weber's Law" for the discrimination of numerosity analogous to the one observed to hold in many sensory domains (Ross 2003; Cantlon and Brannon 2006; Nieder and Merten 2007; Nieder 2013). Moreover, when subjects must report an estimate of the number of dots in a visual array,[7] the standard deviation of the distribution of estimates grows

7. Here, we refer to arrays containing more than five or so dots. As discussed by Jevons (1871) and many subsequent authors, the numerosity of very small arrays can be immediately perceived (without counting) with high accuracy and confidence; the cognitive process used in such cases, termed "subitizing" by Kaufman *et al.* (1949), is quite distinct from the ability to estimate the approximate numerosity of larger arrays, to which the statements in the text refer.

in proportion to the mean estimate, with both the mean and standard deviation being larger when the true number is larger (Izard and Dehaene 2008; Kramer *et al.* 2011); similarly, when subjects are required to produce a particular number of responses (without counting them), the standard deviation of the number produced varies in proportion to the mean number of responses produced—the property of "scalar variability" (Whalen *et al.* 1999; Cordes *et al.* 2001).

All of these observations are consistent with a theory according to which such judgments of numerosity are based on an internal representation that can be represented mathematically by a quantity that is proportional to the logarithm of the numerical value that is being encoded, plus a random error the variance of which is independent of the numerical value that is encoded (van Oeffelen and Vos 1982; Izard and Dehaene 2008). Let the number n be represented by a real number r that is drawn from a distribution

$$r \sim N(\log n, \nu^2), \tag{1.1}$$

where ν is a parameter independent of n. Suppose furthermore that if two stimuli of respective numerosities n_1 and n_2 are presented, their corresponding internal representations r_1, r_2 are independent draws from the corresponding distributions.

Finally, suppose that a subject judges the second array to be more numerous than the first if and only if the internal representations satisfy $r_2 > r_1$. (This is an optimal decision rule, in the sense of maximizing the frequency of correct answers, assuming that any given pair of stimuli are equally likely to be presented in either order.) Then a subject is predicted to respond that array 2 is more numerous with probability

$$\text{Prob}[\text{"2 is more"}] = \Phi\left(\frac{\log(n_2/n_1)}{\sqrt{2}\nu}\right), \tag{1.2}$$

where $\Phi(z)$ is the cumulative distribution function of a standard normal variate z.

Equation (1.2) predicts that "Weber's Law" should be satisfied: the response probability depends only on the ratio n_2/n_1, and not on the absolute numerosity of either array. And indeed, Garcia *et al.* (2018) find that response probabilities are close to being scale-invariant in this sense. The equation also predicts that the z-transformed response probability ($z(p) \equiv \Phi^{-1}(p)$) should be an increasing linear function of $\log(n_2/n_1)$, and hence an approximately linear function of n_2 (for values of n_2 near some fixed value of n_1), with an intercept of zero when $n_2 = n_1$, and a positive slope that decreases for higher values of the reference numerosity n_1; this is exactly what the discrimination data of Krueger (1984) show.[8]

The observed variability of estimates of numerosity is consistent with the same kind of model of noisy coding. Suppose that the subject's estimate $\hat{n}$ of the numerosity of some array must be produced on the basis of the noisy internal representation r hypothesized above. If we approximate the prior distribution from which the true numerosity n is drawn (in a given experimental context) by a log-normal distribution,[9] $\log n \sim N(\mu, \sigma^2)$, then the posterior distribution for n, conditional on an internal representation r drawn from (1.1), will also be log-normal: $\log n | r \sim N(\mu_{\text{post}}(r), \sigma^2_{\text{post}})$. Here $\mu_{\text{post}}(r)$ is an affine function of r, with a slope $0 < \beta < 1$ given by

$$\beta \equiv \frac{\sigma^2}{\sigma^2 + \nu^2}, \tag{1.3}$$

8. See Figure 5 of Krueger (1984), in which the three panels correspond to three successively larger values of n_1, and further discussion in Woodford (2020).

9. We adopt this approximation in order to allow a simple analytical calculation of the posterior distribution, even though in the experiments referred to here, the value of n is actually always an integer. Note that in the application of this model in Section 2, monetary payments are assumed to be positive real numbers rather than integers.

while $\sigma^2_{\text{post}} > 0$ is independent of r (see the Supplementary Appendix for details).

If we hypothesize that the subject's numerosity estimate is optimal, in the sense of minimizing the mean squared estimation error when stimuli are drawn from the assumed prior distribution, then we should expect the subject's estimate to be given by the posterior mean, $\hat{n}(r) = \mathrm{E}[n|r]$. In this case, $\log \hat{n}(r)$ will be an affine function of r, with a slope of β. This result, together with (1.1), implies that conditional on the true numerosity n, the estimate $\hat{n}$ will be log-normally distributed: $\log \hat{n} \sim N(\hat{\mu}(n), \hat{\sigma}^2)$, where $\hat{\mu}(n)$ is an affine function of $\log n$ with slope β, and $\hat{\sigma}^2$ is independent of n. It then follows from the properties of log-normal distributions that

$$\frac{\mathrm{SD}[\hat{n}]}{\mathrm{E}[\hat{n}]} = \sqrt{e^{\hat{\sigma}^2} - 1} > 0,$$

regardless of the true numerosity n. Thus the property of scalar variability is predicted by a model of optimal estimation.

A further implication of a Bayesian model of numerosity estimation is that the average subjective estimate $\mathrm{E}[\hat{n}|n]$ should in general differ from the true numerosity n: subjects' estimates should be *biased.* Specifically, the model just proposed implies a power-law relationship,

$$\mathrm{E}[\hat{n}|n] = An^{\beta} \tag{1.4}$$

for some $A > 0$, where $0 < \beta < 1$ is again defined by (1.3). This implies *over-estimation* of small numerosities (greater than five), but *under-estimation* of larger numerosities, to a progressively greater extent the larger the true numerosity n. This kind of "regressive bias" in subjects' estimates of numerosity is characteristic of all experiments in this area, beginning with the classic study of Kaufman *et al.* (1949). In fact, authors often report that average estimates can be fit reasonably well by a concave power law (or log–log plot), of the kind derived above (Krueger 1972, 1984; Indow and Ida 1977; Kramer *et al.* 2011). The cross-over point, however, at which the bias switches from over-estimation to under-estimation varies across studies; Izard and Dehaene (2008) point out that it depends on the range of numerosities used in the study in question. This is clearly shown by Anobile *et al.* (2012), who find different concave mappings from n to $\mathrm{E}[\hat{n}|n]$ in two experiments using similar methodologies, but different ranges for the true numerosities used in the experiment (1–30 dots in one case, 1–100 dots in the other).

As shown in the Supplementary appendix, this is just what the Bayesian model proposed above would predict: if we vary μ across experiments, holding the other parameters fixed, the cross-over point is predicted to vary in proportion to the variation in the prior mean of n. The Bayesian model also predicts, for a given prior, that increased imprecision in mental coding (a larger value of ν) should result in a lower value of β, and hence a more concave relationship between the true and estimated numerosities; and this is what Anobile *et al.* (2012) find when subjects' cognitive load is increased, by requiring them to perform another perceptual classification task in addition to estimating the number of dots present. Thus, many quantitative features of observed errors in judgments of numerosity are consistent with a model of optimal judgment based on a noisy internal representation of numerosity, and a specific (log-normal) model of the noisy coding of numerical magnitudes in such cases.

1.2. *Symbolically presented numerical information*

The well-documented imprecision in people's perception of visually presented numerical information might seem, however, to be irrelevant for typical laboratory decisions under risk, in which the relevant monetary amounts are described to the decision maker using number symbols.

One might reasonably suppose that symbolically presented numbers are generally understood precisely by the hearer; and to the extent that perceptual errors do occur, they should not generally be expected to conform to Weber's Law, as in the case of sensory magnitudes. (If it were a simple matter of sometimes mis-hearing numbers stated by an experimenter, one might expect that $34.13 could more easily be mistaken for $44.13 than for $34.89.)

Nonetheless, there is a good deal of evidence suggesting that even when numerical quantities are presented using symbols such as Arabic numerals, the semantic content of the symbol is represented in the brain in a way that is similar to the way in which magnitudes are represented—involving imprecision, just as with the representation of physical magnitudes, and with similar quantities represented in similar ways, so that nearby numerical magnitudes are more likely to be confused with one another (Dehaene 2011). This is not the *only* way in which numerical information is understood to be represented in the brain; according to the well-known "triple-code model" of Dehaene (1992), numbers are represented in three different ways (three "codes"), in circuits located in different regions of the brain, each with a distinct function. An "Arabic code" is used for explicit multi-digit arithmetic calculations, while simple verbal counting and retrieval of memorized facts of arithmetic are instead executed via a "verbal code."

Yet a third code, the "analog magnitude code," is drawn upon in tasks involving number comparisons and approximation. This is thought to be a "semantic" representation of the size of the quantity represented by a given number—"the abstract quantity meaning of numbers rather than the numerical symbols themselves" (Dehaene *et al.* 2003, p. 492)—and to be independent of the symbolic form in which the number is presented. Scalp electroencephalography (EEG) recordings while subjects process information presented in the form of Arabic numerals also indicate that the neural patterns evoked by particular numbers vary continuously with numerical distance, so that (*e.g.*) the neural signals for "3" are more similar to those for "4" than to those for "5" (Spitzer *et al.* 2017; Luyckx *et al.* 2018; Teichmann *et al.* 2018).

The existence of an approximate semantic representation of numerical quantities, even when numbers are presented symbolically, can also be inferred behaviourally from the ability of patients with brain injuries that prevent them from performing even simple arithmetic (using the exact facts of arithmetic learned in school) to nonetheless make fairly accurate approximate judgments (Dehaene and Cohen 1991). In normal adult humans, this approximate "number sense" seems also to be drawn upon when number comparisons are made very quickly, or when previously presented numerical information that has not been precisely memorized must be recalled (Moyer and Landauer 1967; Dehaene *et al.* 1990).

Moreover, there is evidence that the mental representation of numerical information used for approximate calculations involves the same kind of logarithmic compression as in the case of non-symbolic numerical information, even when the numerical magnitudes have originally been presented symbolically. Moyer and Landauer (1967), Buckley and Gillman (1974), and Banks *et al.* (1976) find that the reaction time required to judge which of two numbers (presented as numerals) is larger varies with the distance between the numbers on a compressed, non-linear scale—a logarithmic scale, as assumed in the model of the coding of numerosity sketched above, or something similar—rather than the linear (arithmetic) distance between them. Further evidence suggesting that such judgments are based on imprecise analog representations of the numbers presented comes from the finding of Frydman and Jin (2019) that the distance between numbers required for their relative size to be correctly judged with a given probability shrinks when the range of variation in the numbers presented in the experiment is smaller; such an effect is difficult to explain if errors are attributed to mistakes in processing the presented number symbols.[10]

10. Their result can instead be explained by the model of logarithmic coding presented in the previous section, under a small extension of the model discussed in the Supplementary Appendix.

In an even more telling example for our purposes, Dehaene and Marques (2002) showed that in a task where people had to estimate the prices of products, the estimates produced exhibited the property of scalar variability, just as with estimates of the numerosity of a visual display. This was found to be the case, even though both the original information people had received about prices and the responses they produced involved symbolic representations of numbers. Evidently, an approximate analog representation of the prices remained available in memory, though the precise symbolic representation of the prices could no longer be accessed.

Not only is there evidence for the existence of an approximate semantic representation of numerical information that is presented symbolically; it seems likely that this "analog magnitude code" is the *same* representation of number that is used when numbers are presented non-symbolically. The region in the intraparietal sulcus that is thought to be the locus of the analogue magnitude code seems to be activated by the presentation of numerical stimuli, regardless of the format in which the information is presented: written words or Arabic numerals, visual or auditory presentations, symbolic or non-symbolic (Piazza *et al.* 2004; Brannon 2006). If this is true, it means that we should expect the quantitative model of imprecise internal representations that explains the perception of numerosity, a context in which the statistical structure of errors has been documented in more detail, to also apply to the imprecise internal representations that are drawn upon when fast, approximate judgments are made about symbolically presented numerical information. We shall explore the implications of this hypothesis for risky choice.

More specifically, our hypothesis is that when people must decide whether a risky prospect (offering either of two possible monetary amounts as the outcome) is worth more or less than another monetary amount that could be obtained with certainty, they can make a quick, intuitive judgment about the relative value of the two options using the same mental faculty as is involved in making a quick estimate (without explicit use of precise arithmetic calculations) as to whether the sum of two numbers is greater or less than some other number.

If this is approached as an approximate judgment rather than an exact calculation (as will often be the case, even with numerate subjects), such a judgment is made on the basis of mental representations of the monetary amounts that are approximate and analogue, rather than exact and symbolic; and these representations involve a random location of the amount on a logarithmically compressed "mental number line." The randomness of the internal representation of the numerical quantity (or perhaps, of its value to the decision maker) then provides an explanation for the randomness in laboratory decisions as to whether to accept simple gambles; and as we show below, the logarithmic compression provides an explanation for subjects' apparent risk aversion, even in the case of gambles for fairly small stakes.

Note that we do not assume that all decisions involving money are made in this way. If someone is asked to choose between \$20 and \$22, either of which can be obtained with certainty, we do not expect that they will sometimes choose the \$20, because of noise in their subjective sense of the size of these two magnitudes. The question whether \$20 is greater or smaller than \$22 can instead be answered reliably (by anyone who remembers how to count), using the "verbal code" hypothesized by Dehaene (1992) to represent the numbers, rather than the "analogue magnitude code."

Likewise, we do not deny that numerate adults, if they take sufficient care (and consciously recognize the problem facing them as having the mathematical structure of a type of arithmetic problem), are capable of exact calculations of averages or expected values that would not introduce the kind of random error modeled in the next section. Nonetheless, we hypothesize that questions about small gambles in laboratory settings (even when incentivized) are often answered on the basis of an intuitive judgment based on approximate analog representations of the quantities involved. Note also that our hypothesis does not depend on an assumption that numerical quantities are mis-perceived at the time that the problem is described to the subject; our model of lottery choice is perfectly consistent with the subject being able to repeat back to the experimenter the quantities that he has been told are at stake. But even when the subject knows exactly what the numbers are (*i.e.* has access to an exact description of them using the "verbal code"), if the decision problem is not trivial to answer on the basis of these numbers, we suppose that he may resort to an approximate judgment on the basis of the imprecise semantic representation of the numbers, present in the brain at the same time.

While our results here cannot prove this, we suspect that many economic decisions in everyday life are also made on the basis of approximate calculations using imprecise semantic representations of numerical magnitudes. The situation in typical laboratory experiments studying choice under risk is actually the one that is most favourable to the use of explicit mental arithmetic: the possible payoffs are completely enumerated and explicitly stated, and the associated probabilities are explicitly stated as well. If, as our results suggest, choices can be based on approximate calculations of the kind that we model even in such a simple and artificial setting, it seems even more likely that cognitive mechanisms of this kind are employed in real situations where the relevant data are only estimates to begin with.

2. A MODEL OF NOISY CODING AND RISKY CHOICE

We now consider the implications of a model of noisy internal representation of numerical magnitudes for choices between simple lotteries. We assume a situation in which a subject is presented with a choice between two options: receiving a monetary amount $C > 0$ with certainty, or receiving the outcome of a lottery, in which she will have a probability $0 < p < 1$ of receiving a monetary amount $X > 0$. We wish to consider how decisions should be made if they must be based on imprecise internal representations of the monetary amounts rather than their exact values.

We hypothesize that the subject's decision rule is optimal, in the sense of maximizing the expected value of $U(W)$, subject to the constraint that the decision must be based on an imprecise representation $\mathbf{r}$ of the problem, rather than the true data. Here, W is the subject's final wealth at the end of the experiment, and $U(W)$ is an indirect utility function, assumed to be smooth and strictly increasing, indicating the (correctly assessed) expected value to the subject of a given wealth. Note that our conceptualization of the subject's objective (from the standpoint of which the decision rule can be said to be optimal) involves no "narrow bracketing" of the gains from a particular decision: it is assumed that only final wealth W matters, and not the sequence of gains and losses by which it is obtained. The expected value is defined with respect to some prior probability distribution over possible decision situations (here, possible values of X and C that might be offered).

Let W^a be the random final wealth if option a is chosen. If we consider only gambles for small amounts of money, we can use the Taylor approximation $U(W^a) \approx U(W_0) + U'(W_0) \cdot \Delta W^a$, where W_0 is the subject's wealth apart from any gain or loss from the experiment, ΔW^a is the random monetary amount gained in the experiment if option a is chosen, and $U'(W_0)$ is positive for all possible values of W_0. If we assume furthermore that the subject's information about W_0 is coded by some internal representation r_0, with a distribution that is independent of the details of the gains offered by the decision problem, while the quantities X and C have internal representations r_x and r_c respectively, that are distributed independently of W_0, then

$$\mathrm{E}[U(W^a)|\mathbf{r}] \approx \mathrm{E}[U(W_0)|r_0] + \mathrm{E}[U'(W_0)|r_0] \cdot \mathrm{E}[\Delta W^a|r_x, r_c]$$

will be an increasing function of $\mathrm{E}[\Delta W^a|r_x, r_c]$, regardless of the value of r_0.

It follows that, as long as stakes are small enough, an optimal decision rule is one that chooses the action a for which the value of $\mathrm{E}[\Delta W^a|r_x, r_c]$ is larger; we therefore consider the hypothesis that decisions are optimal in this sense. Note that our theory's predictions are thus consistent with "narrow bracketing": the choice between two risky prospects is predicted to depend only on the distributions of possible net gains associated with those prospects, and not on the level of wealth W_0 that the subject has from other sources. But for us this is a *conclusion* (a property of optimal decision rules) rather than a separate *assumption.* Note also that while we do not deny the reasonableness of assuming that the function $U(W)$ should involve diminishing marginal utility of wealth (in the case of sufficiently large changes in wealth), the degree of curvature of the function $U(W)$ plays no role in our predictions. Thus small-stakes risk aversion is not attributed to non-linear utility of income or wealth in our theory.

In line with the evidence discussed in the previous section regarding internal representations of numerical magnitudes, we assume more specifically that the representations r_x and r_c are each a random draw from a probability distribution of possible representations, with distributions

$$r_x \sim N(\log X, \nu^2), \qquad r_c \sim N(\log C, \nu^2). \tag{2.1}$$

Here, $\nu > 0$ is a parameter that measures the degree of imprecision of the internal representation of such quantities (assumed to be the same regardless of the monetary amount that is represented); we further assume that r_x and r_c are distributed independently of one another. We treat the parameter p as known (it does not vary across trials in the experiment described below), so that the decision rule can (and indeed should) depend on this parameter as well.[11]

As in the model of numerosity perception presented in Section 1.2, these representations do not themselves constitute perceived values of the monetary amounts; instead, the internal representations must be "decoded" in order to provide a basis for decision, in the case of a given decision problem. The optimal decision in the case of a pair of mental representations $\mathbf{r} = (r_x, r_c)$ depends not only on the specification (2.1) of the noisy coding, conditional on the true magnitudes, but also on the relative ex ante likelihood of different possible decision situations, which we specify by a prior probability distribution over possible values of (X, C). We can then consider the optimal decision rule from the standpoint of Bayesian decision theory. It is easily seen that $\mathrm{E}[\Delta W^a | r_x, r_c]$ is maximized by a rule under which the risky lottery is chosen if and only if

$$p \cdot \mathrm{E}[X|r_x] > \mathrm{E}[C|r_c], \tag{2.2}$$

which is to say if and only if the expected payoff from the risky lottery exceeds the expected value of the certain payoff.[12]

The implications of our logarithmic model of noisy coding are simplest to calculate if (as in the model of numerosity estimation) we assume a log-normal prior distribution for possible monetary quantities. To reduce the number of free parameters in our model, we assume that under the prior X and C are assumed to be independently distributed, and furthermore that the prior distributions for both X and C are the same (some ex ante distribution for possible payments that one may be offered in a laboratory experiment). It is then necessary only to specify the parameters of this common prior:

$$\log X, \log C \sim N(\mu, \sigma^2). \tag{2.3}$$

Under the assumption of a common prior for both quantities, the common prior mean μ does not affect our quantitative predictions about choice behaviour; instead, the value of σ does matter, as

11. See Section 4.1 for discussion of an extension of the model in which p is also imprecisely represented.

12. Note that while the payoff C is certain, rather than random, once one knows the decision situation (which specifies the value of C), it is a random variable ex ante (assuming that many different possible values of C might be offered), and continues to be random even conditioning on a subjective representation of the current decision situation, assuming that mental representations are noisy as assumed here.

this influences the ex ante likelihood of X being sufficiently large relative to C for the gamble to be worth taking. The model thus has two free parameters, to be estimated from subjects' behaviour: σ, indicating the degree of ex ante uncertainty about what the payoffs might be, and ν, indicating the degree of imprecision in the coding of information that is presented about those payoffs on a particular trial.

2.1. *Predicted frequency of acceptance of a gamble*

Under this assumption about the prior, the posterior distributions for both X and C are log-normal, as in the model of numerosity estimation in the previous section. It follows that the posterior means of these variables are given by

$$\mathrm{E}[X|\mathbf{r}] = e^{\alpha+\beta r_x}, \qquad \mathrm{E}[C|\mathbf{r}] = e^{\alpha+\beta r_c},$$

with β is again defined by (1.3). (The details are explained in the Supplementary Appendix.) Taking the logarithm of both sides of (2.2), we see that this condition will be satisfied if and only if

$$\log p + \beta r_x > \beta r_c,$$

which is to say, if and only if the internal representation satisfies

$$r_x - r_c > \beta^{-1} \log p^{-1}. \tag{2.4}$$

Under our hypothesis about the mental coding, r_x and r_c are independently distributed normal random variables (conditional on the true decision situation), so that

$$r_x - r_c \sim N(\log X/C, 2\nu^2).$$

It follows that the probability of (2.4) holding, and the risky gamble being chosen, is given by

$$\text{Prob}[\text{accept risky}|X,C] = \Phi\left(\frac{\log X/C - \beta^{-1}\log p^{-1}}{\sqrt{2}\nu}\right). \tag{2.5}$$

Equation (2.5) is the behavioural prediction of our model. It implies that choice in a problem of this kind should be stochastic, as is typically observed. Furthermore, it implies that across a set of gambles in which the values of p and C are the same in each case, but the value of X varies, the probability of acceptance should be a continuously increasing function of X. This is in fact what one sees in Figure 1, which plots data from Mosteller and Nogee (1951).[13] The figure plots the responses of one of their subjects to a series of questions of the type considered here. In each case, the subject was offered a choice of the form: are you willing to pay five cents for a gamble that will pay an amount X with probability 1/2, and zero with probability 1/2? The figure shows the fraction of trials on which the subject accepted the gamble, in the case of each of several different values of X. The authors used this curve to infer a value of X for which the subjects would be indifferent between accepting and rejecting the gamble, and then proposed to use this value of X to identify a point on the subject's utility function.

13. This study is of particular interest for our purposes because the authors use a method intended to elicit repeated, independent decisions about exactly the same pair of gambles at different points during the same study, as is common in psychometric studies of imprecision in perception. We follow the same method in our own experiment, reported in the next section.

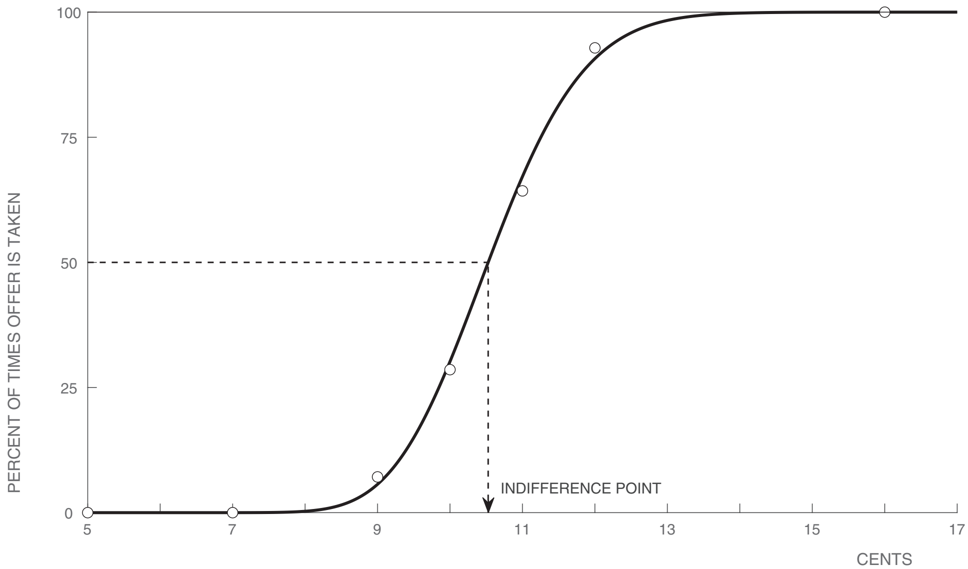

FIGURE 1

Theoretically predicted probability of acceptance of a simple gamble as a function of the value X when the gamble pays off, for parameter values explained in the text. Circles represent the data from Mosteller and Nogee (1951).

Figure 1 plots the data from Mosteller and Nogee (1951), together with a solid curve that graphs the predicted relationship (2.5), in the case that $\sigma = 0.26$ and $\nu = 0.07$. Note that these values allow a reasonably close fit to the choice frequencies plotted in the figure from Mosteller and Nogee.

Moreover, the parameter values required to fit the data are fairly reasonable ones. The value $\nu = 0.07$ for the so-called "Weber fraction" is less than half the value indicated by human performance in comparisons of the numerosity of different fields of dots (Dehaene 2008, p. 540); on the other hand, Dehaene (2008, p. 552) argues that one should expect the Weber fraction to be smaller in the case of numerical information that is presented symbolically (as in the experiment of Mosteller and Nogee) rather than non-symbolically (as in the numerosity comparison experiments).[14] Hence this value of ν is not an implausible degree of noise to assume in the mental representations of numerical magnitudes used in approximate calculations. The value of σ for the degree of dispersion of the prior over possible monetary rewards is also plausible. In fact, the distribution of values of X used in the trials reported in the figure is one in which the standard deviation of $\log X$ is 0.34, so that the implicit prior attributed to the subject by our model is of at least the right order of magnitude.

2.2. *Explaining the Rabin paradox*

Our model explains not only the randomness of the subject's choices, but also her apparent risk aversion, in the sense that the indifference point (a value of X around 10.7 cents in Figure 1) corresponds to a gamble that is better than a fair bet. This is a general prediction of the model,

14. Garcia *et al.* (2018) have the same subjects do two versions of such a task, one in which the monetary amounts are presented symbolically (as in the experiment of Mosteller and Nogee) and one in which the amounts X and C are presented as visual arrays of euro coins. They find similar stochastic choice curves in both cases, but the implied value of ν is larger when the amounts are shown visually.

since the indifference point is predicted to be at $X/C = (1/p)^{\beta^{-1}} > 1/p$, where the latter quantity would correspond to a fair bet. The model predicts risk neutrality (indifference when $X/C = 1/p$) only in the case that $\beta = 1$, which according to (1.3) can occur only in the limiting cases in which $\nu = 0$ (perfect precision of the mental representation of numerical magnitudes), or σ is unboundedly large (radical uncertainty about the value of the payoff that may be offered, which is unlikely in most contexts).

The model furthermore explains the Rabin (2000) paradox: the fact that the compensation required for risk does not become negligible in the case of small bets. According to EUM, the value of X required for indifference in a decision problem of the kind considered above should be implicitly defined by the equation

$$pU(W_0+X)+(1-p)U(W_0) = U(W_0+C).$$

For any increasing, twice continuously differentiable utility function $U(W)$ with $U'' < 0$, if $0 < p < 1$, this condition implicitly defines a solution $X(C;p)$ with the property that $pX(C;p)/C > 1$ for all $C > 0$, implying risk aversion. However, as C is made small, $pX(C;p)/C$ necessarily approaches 1. Hence the ratio pX/C required for indifference exceeds 1 (the case of a fair bet) only by an amount that becomes arbitrarily small in the case of a small enough bet. It is not possible for the required size of pX to exceed the certain payoff even by 7% (as in the case shown in Figure 1), in the case of a very small value certain payoff, unless the coefficient of absolute risk aversion $(-U''/U')$ is very large—which would in turn imply an implausible degree of caution with regard to large bets.

In our model, instead, the ratio pX/C required for indifference should equal $\Lambda \equiv p^{-(\beta^{-1}-1)}$, which is greater than 1 (except in the limiting cases mentioned above) by the same amount, regardless of the size of the gamble. As discussed above, the degree of imprecision in mental representations required for Λ to be on the order of 1.07 is one that is quite consistent with other evidence. Hence, the degree of risk aversion indicated by the choices in Figure 1 is wholly consistent with a model that would predict only a modest degree of risk aversion in the case of gambles involving thousands of dollars.

It is also worth noting that our explanation for apparent risk aversion in decisions about small gambles does not rely on loss aversion, like the explanation proposed by Rabin and Thaler (2001). Our model of the mental representation of prospective gains assumes that the coding and decoding of the risky payoff X are independent of the value of C, so that small increases in X above C do not have a materially different effect than small decreases of X below C.

Instead, in our theory the EUM result that the compensation for risk must become negligible in the case of small enough gambles fails for a different reason. Condition (2.4) implies that the risky gamble is chosen more often than not if and only if $p \cdot m(X) > m(C)$, where $m(\cdot)$ is a power-law function of a kind that also appears in (1.4). It is *as if* the decision maker assigned a non-linear utility $m(\Delta W^a)$ to the wealth increment ΔW^a. Our model of optimal decision on the basis of noisy internal representations explains why the ratio $m(X)/m(C)$ is in general not approximately equal to X/C even in the case that X and C are both small.

3. AN EXPERIMENTAL TEST

A notable feature of the behavioural equation (2.5) is that it predicts that subjects' choice frequencies should be *scale-invariant*, at least in the case of all small enough gambles: multiplying both X and C by an arbitrary common factor should not change the probability of the risky gamble being chosen. This feature of the model makes it easy to see that the Rabin paradox is not problematic for our model. In order to test this predictions of our model, we conducted an experiment of our own, in which we varied the magnitudes of both X and C. We recruited 20 subjects from the student population at Columbia University, each of whom was presented with a sequence of several 100 trials. Each individual trial presented the subject with a choice between a certain monetary amount C and a probability p of receiving a monetary amount X.[15]

The probability p of the non-zero outcome under the lottery was 0.58 on all of our trials, as we were interested in exploring the effects of variations in the magnitudes of the monetary payments, rather than variations in the probability of rewards, in order to test our model of the mental coding of monetary amounts. Maintaining a fixed value of p on all trials, rather than requiring the subject to pay attention to the new value of p associated with each trial, also made it more plausible to assume (as in the model above) that the value of p should be known precisely, rather than having to be inferred from an imprecisely coded observation on each occasion.

We chose a probability of 0.58, rather than a round number (such as one-half, as in the Mosteller and Nogee experiment discussed above), in order not to encourage our subjects to approach the problem as an arithmetic problem that they should be able to solve exactly, on the basis of representations of the monetary amounts using the "Arabic code" rather than the "analog magnitude code," in the terminology of Dehaene (1992). We expect Columbia students to be able to solve simple arithmetic problems using methods of exact mental calculation that are unrelated to the kind of approximate judgments about numerical magnitudes with which our theory is concerned but did not want to test this in our experiment. We chose dollar magnitudes for C and X on all trials that were not round numbers, either, for the same reason.

The value of the certain payoff C varied across trials, taking on the values \$5.55, \$7.85, \$11.10, \$15.70, \$22.20, or \$31.40. (Note that these values represent a geometric series, with each successive amount $\sqrt{2}$ times as large as the previous one.) The non-zero payoff X possible under the lottery option was equal to C multiplied by a factor $2^{m/4}$, where m took an integer value between 0 and 8. There were thus only a finite number of decision situations (defined by the values of C and X) that ever appeared, and each was presented to the subject several times over the course of a session. This allowed us to check whether a subject gave consistent answers when presented repeatedly with the same decision, and to compute the probability of acceptance of the risky gamble in each case, as in the experiment of Mosteller and Nogee. The order in which the various combinations of C and X were presented was randomized, in order to encourage the subject to treat each decision as an independent problem, with the values of both C and X needed to be coded and encoded afresh, and with no expectations about these values other than a prior distribution that could be assumed to be the same on each trial.

Our experimental procedure thus differed from ones often used in decision-theory experiments, where care is taken to present a sequence of choices in a systematic order, so as to encourage the subject to express a single consistent preference ordering. We were instead interested in observing the randomization that, according to our theory, should occur across a series of genuinely independent reconsiderations of a given decision problem; and we were concerned to simplify the context for each decision by eliminating any obvious reason for the data of one problem to be informative about the next.

We also chose a set of possible decision problems with the property that each value of X could be matched with the same geometric series of values for C, and vice versa, so that on each trial it was necessary to observe the values of both C and X in order to recognize the problem, and neither value provided much information about the other (as assumed in our theoretical model). At the same time, we ensured that the ratio X/C, on which the probability of choosing the lottery should depend according to our model, always took on the same finite set of values for each value

15. The experimental design is discussed further in the Supplementary Appendix. Our procedures were approved by the Columbia University Institutional Review Board, under protocol IRB-AAAQ2255.

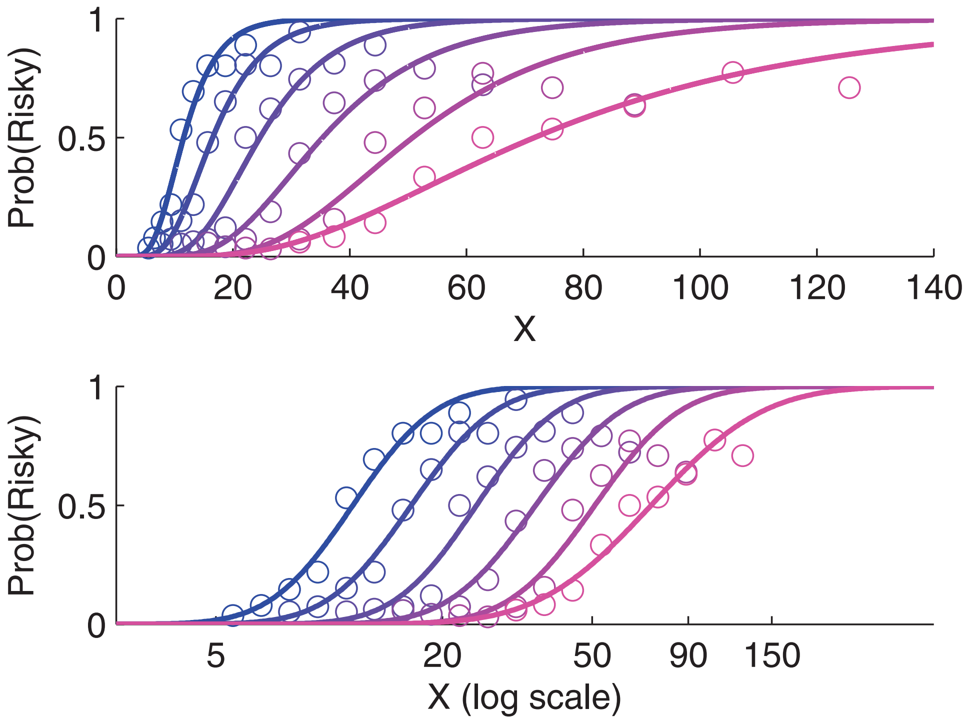

FIGURE 2

The probability of choosing the risky lottery, plotted as a function of the risky payoff X (data pooled from all 20 subjects). (a) The probability plotted as a function of X, for each of the different values of C (indicated by darkness of lines). (b) The same figure, but plotted against $\log X$ for each value of C.

of C. This allowed us to test whether the probability of choosing the lottery would be the same when the same value of X/C recurred with different absolute magnitudes for X and C.

3.1. *Testing scale-invariance*

Figure 2 shows how the frequency with which our subjects chose the risky lottery varied with the monetary amount X that was offered in the event that the gamble paid off, for each of the five different values of C.[16] (For this first analysis, we pool the data from all 20 subjects.) Each data point in the figure (shown by a circle) corresponds to a particular combination (C,X).

In the first panel, the horizontal axis indicates the value of X, while the vertical axis indicates the frequency of choosing the risky lottery on trials of that kind [Prob(Risky)]. The different values of C are indicated by different colours of circles, with the darker circles corresponding to the lower values of C, and the lighter circles the higher values. (The six successively higher values of C are the ones listed above.) We also fit a sigmoid curve to the points corresponding to each of the different values of C, where the colour of the curve again identifies the value of C. Each curve has an equation of the form

$$\text{Prob(Risky)} = \Phi(\delta_C + \gamma_C \log X), \tag{3.1}$$

16. The data used to produce this and all subsequent figures, as well as the results reported in the tables, can be found in Khaw *et al.* (2020).

where $\Phi(z)$ is again the CDF of the standard normal distribution, and the coefficients (δ_C, γ_C) are estimated separately for each value of C so as to maximize the likelihood of the data corresponding to that value of C. Note that for each value of C, we obtain a sigmoid curve similar to the one in Figure 1, though the fit is less perfect (at least partly because here, unlike in Figure 1, we are pooling the data from 20 different subjects).

The similarity of the curves obtained for different values of C can be seen more clearly if we plot them as a function of $\log X$, rather than on a scale that is linear in X, as shown in the second panel of Figure 2. (The colour coding of the curves corresponding to different values of C is again the same.) The individual curves now resemble horizontal shifts of one another. The elasticity γ_C is similar for each of the values of C (with the exception of the highest value, $C = \$31.40$), and the value of $\log X$ required for indifference increases by a similar amount each time C is multiplied by another factor of $\sqrt{2}$.

These observations are exactly what we should expect, according to our logarithmic coding model. Condition (2.5) implies that a relationship of the form

$$\text{Prob(Risky)} = \Phi(\delta + \gamma \log(X/C)) \tag{3.2}$$

should hold for all values of C, meaning that in equation (3.1), γ_C should be the same for each value of C, and that the value of $\log X$ required for indifference should equal a constant plus $\log C$. We can see more clearly the extent to which these precise predictions hold by plotting the curves in Figure 2(b) as functions of $\log(X/C)$, rather than as functions of $\log X$; this is done in the first panel of Figure 3. The six different curves come close to falling on top of one another, as predicted by the model (although, again, the curve for $C = \$31.40$ is somewhat out of line with the others). If we instead simply estimate parameters (δ, γ) to maximize the likelihood of the pooled data under the model (3.2), we obtain the single choice curve shown in the second panel of Figure 3. This fits the data for the different values of X/C slightly worse than the individual choice curves shown in the previous panel but not by much. (The maximum-likelihood parameter estimates for the different choice curves, and the associated likelihoods, are reported in the Supplementary Appendix.)

We can consider quantitatively the extent to which our data are more consistent with the more flexible model (3.1) than with the more restrictive predictions of (3.2), in two different ways. First, we consider the *in-sample* fit of the two models by selecting a subset of our observations (the "calibration dataset"), and find the parameter estimates for each model that maximize the likelihood of this dataset. The column labelled $LL^{\text{calibration}}$ in Table 1 reports the average maximized value of the log-likelihood of the data in the calibration dataset, when it is chosen in each of four different ways. We use a "four-fold cross-validation" approach, in which the complete dataset is divided into four parts, each containing exactly 1/4 of the observations, and the model parameters are estimated four different times; each time, one of the four parts of the data is held out to be the "validation dataset," and the other three parts are used as the "calibration dataset." Thus in the overall exercise, each observation is used equally many times as part of the calibration dataset as every other. The figures in Table 1 report the average values of the statistics obtained in the four different estimations. With regard to in-sample fit, of course, $LL^{\text{calibration}}$ is higher for the more flexible model, since (3.2) is nested within this class of models as a special case.

A more relevant comparison between the in-sample fits of the two models is given by the Bayes information criterion (BIC) statistic, also reported in the table for each model, which penalizes the use of additional free parameters. This is defined as $\text{BIC} \equiv -2LL + k \log N_{\text{obs}}$, where k is the number of free parameters (adjusted to maximize the likelihood) for a given model, and N_{obs} is

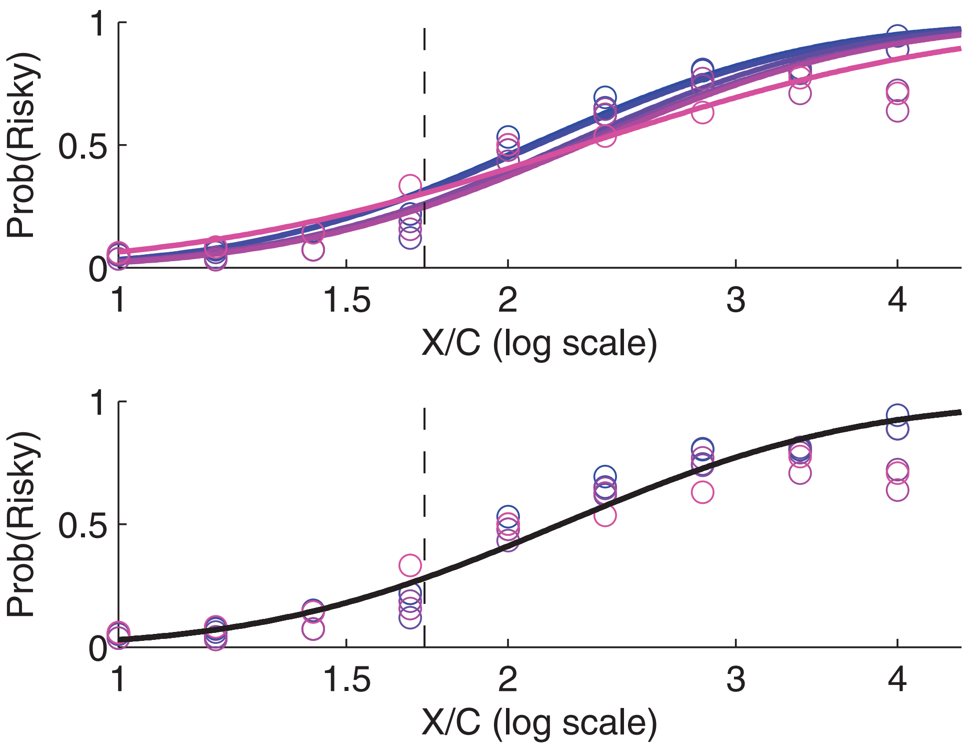

FIGURE 3

The same data as in Figure 2, now plotted as a function of $\log X/C$. (a) A separate choice curve estimated for each value of C, as in Figure 2. (b) A single choice curve, with parameters estimated to maximize the likelihood of the pooled data.

TABLE 1

In-sample and out-of-sample measures of goodness of fit compared for the scale-invariant model (our logarithmic coding model) and an unrestricted statistical model in which a separate choice curve is estimated for each value of C. In the top panel, each model is fit to the pooled data from all 20 subjects. In the bottom panel, separate model parameters are fit to the data for each subject. (See text for further explanation.)

Model	$LL^{\text{calibration}}$	BIC	$LL^{\text{validation}}$	log K
		Pooled data		
Scale-invariant	−2,812.7	5,642.9	−940.0	0.0
Unrestricted	−2,794.9	5,672.5	−938.1	12.9
		Pooled data		
Scale-invariant	−1,853.4	3,932.6	−677.5	0.0
Unrestricted	−1,556.1	3,962.2	−2,214.8	1,552.1

the number of observations in the calibration dataset.[17] The data provide more evidence in favour of the model with the lower BIC statistic. In particular, for any two models $\mathcal{M}_1$ and $\mathcal{M}_2$, the *Bayes factor K* defined by

$$\log K_1 = \frac{1}{2}[\text{BIC}(\mathcal{M}_2) - \text{BIC}(\mathcal{M}_1)]$$

17. Again, the BIC statistics reported in the table are actually the average values of the four BIC statistics obtained for the four different choices of calibration dataset. And as elsewhere in the article, "log" refers to the natural logarithm.

is the multiplicative factor by which the relative posterior probability that $\mathcal{M}_1$ rather than $\mathcal{M}_2$ is the correct model of the data is increased by the observations in the calibration dataset. (See, for example, Burnham and Anderson (2002), p. 303.)

We can also compare the *out-of-sample* fit of the two models, by reserving some of our observations (the "validation dataset"), and not using them to estimate the model parameters. The column labelled $LL^{\text{validation}}$ in Table 1 then reports the average log-likelihood of the data in the validation dataset under each model, when the parameter values are used that were estimated using the calibration dataset. If we update the posterior probabilities that the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ are correct after observing the validation dataset as well, we obtain a composite Bayes factor $K = K_1 \cdot K_2$, where

$$\log K_2 = LL^{\text{validation}}(\mathcal{M}_1) - LL^{\text{validation}}(\mathcal{M}_2)$$

by Bayes' Rule. The logarithm of the composite Bayes factor K is reported in the final column of the table, as an overall summary of the degree to which the data provide support for each model, averaged across four different ways of choosing the validation dataset. (In each case, $\mathcal{M}_1$ is the scale-invariant model, while $\mathcal{M}_2$ is the alternative model considered on that line of the table; thus values $K > 1$ indicate the degree to which the data provide more support for the scale-invariant model than for the alternative.)

In Table 1, we compare two models: our scale-invariant model (3.2) and the unrestricted alternative in which a separate probit model (3.1) is estimated for each of the six values of C, as in Figure 2.[18] In the case of the scale-invariant model, N_{obs} is the total number of observations in the calibration dataset, pooling the data for all six values of C, and there are $k = 2$ free parameters in the single model fit to all of these data. In the case of the unrestricted model, a separate probit model (each with $k = 2$ free parameters) is estimated for each value of C, and a BIC statistic is computed for that model (where N_{obs} is the number of observations in the calibration dataset with that value of C); the BIC reported in the "Unrestricted" row of the table is then the sum of the BIC statistics for these six independent probit models, just as $LL^{\text{calibration}}$ is the sum of the log likelihoods for the six models. (More precisely, the table shows the average value of this sum of BIC statistics, for each of the four different "folds" of the data.) In the top panel of the table, the two models are compared when a common set of parameters is used to fit the pooled data from all 20 subjects, as in Figures 2 and 3. In the lower panel, instead, individual model parameters are estimated for each subject, and the statistics reported are sums over all subjects of the corresponding model fit statistics for each subject.

Whether we assume a common set of parameters or subject-specific parameters, we see that the BIC statistic is lower for the scale-invariant model. This means that while the unrestricted model achieves a higher likelihood (necessarily), the data are not fit enough better to justify the use of so many additional free parameters; thus based on the calibration dataset alone, we would have a Bayes factor $K_1 > 1$, meaning an increase in the relative posterior probability of the scale-invariant model (compared to whatever relative probability was assigned to that model in one's prior). When we then consider out-of-sample fit of the two models, if we assume a common set of parameters for all 20 subjects, the out-of-sample fit is slightly better for the unrestricted model. However, the fit is only modestly better, and when one takes into account both the in-sample

18. Note that the scale-invariant model and unrestricted alternative referred to in Table 1 do not correspond precisely to the predictions shown in Figures 2 and 3, since in Figures 2 and 3 the parameters of both models are fit to our entire dataset, while in Table 1 the parameters are estimated using only a subset of the data.

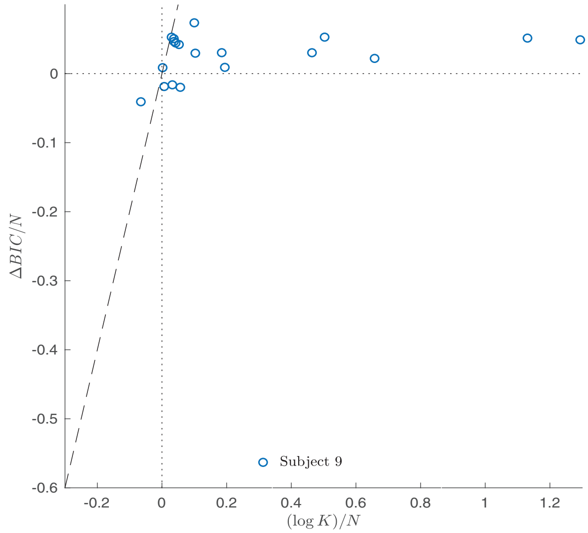

FIGURE 4

In-sample and out-of-sample model comparison statistics, for each of the 20 individual subjects, when separate parameters are estimated for each subject. (See explanation in text.)

and out-of-sample fit of the two models, we obtain an overall Bayes factor $K > 400,000$, greatly increasing the relatively probability of the scale-invariant model.[19]

If we instead fit separate parameters for each subject, then as shown in the bottom panel of Table 1, the aggregate evidence provides more support for the scale-invariant model both in-sample and out-of-sample. In this case, the overall Bayes factor is greater than 10^{674}. Thus, if we assume that either the scale-invariant model or the more flexible alternative must be the correct model for all subjects (though the parameters may differ across subjects), the evidence is overwhelming in favour of the scale-invariance hypothesis.

In fact, the scale-invariant model fits reasonably well for most of the subjects considered individually. Figure 4 shows a scatter plot of the values of the BIC difference, and the overall Bayes factor K, for each individual subject, when separate choice curves are estimated for each subject. Here, each open dot corresponds to one subject. The vertical axis plots the amount by which the BIC statistic for the unrestricted model is greater than the one for the scale-invariant model (ΔBIC), divided by N, the number of trials for that subject, in order to obtain a measure that is more comparable across subjects. The horizontal axis plots the value of $\log K$, again divided by N. (In both cases, the values plotted for each subject are the average of the values obtained for the four different "folds" of the data.) The dashed line identifies points at which $\log K = (1/2)\Delta\text{BIC}$, which is to say, points at which there is no difference in $LL^{\text{valuation}}$ between the two models. Points to the right of the dashed line are thus those for which $LL^{\text{valuation}}$ is higher for the scale-invariant model than for the unrestricted model. We see that the overall Bayes factor favours the scale-invariant model for all but one of the subjects (subject 14). Moreover, the scale-invariant model fits better (or approximately as well) out-of-sample in the case of all of those nine subjects;

19. Moreover, the slight inferiority of the scale-invariant model with regard to out-of-sample fit is due primarily to the data for a single subject (subject 9), whose choice curves do not satisfy scale-invariance, as shown in the Supplementary Appendix. If we fit a single set of parameters to the pooled data for all subjects except subject 9, the scale-invariant model fits better both in-sample and out-of-sample.

while it is favoured in-sample by the BIC criterion for 15 out of 20 (only fitting substantially worse in-sample for subject 9).

Garcia *et al.* (2018) and Frydman and Jin (2019) repeat versions of our experiment, and similarly find near scale-invariance of the choice curves associated with different values of C, though they do not report statistical tests of scale invariance like those above. See, for example, the bottom panel of Figure B2 in Frydman and Jin.

Despite this degree of support for our model's prediction, our data are nonetheless not perfectly scale-invariant. We see in Figure 3 that the estimated choice curve (using pooled data) in the case $C = \$31.40$ is not a perfect horizontal translation of the others, but instead is somewhat flatter.[20] This may indicate inaccuracy of the assumption of a log-normal prior (2.3), used in our theoretical calculations above for convenience. Under the assumption of a log-normal prior, $\log \mathrm{E}[X|r_x]$ is a linearly increasing function of r_x, with constant slope β. But if people instead form correct inferences based on a prior under which monetary payments greater than \$50 are less likely than a log-normal prior would allow, then $\log \mathrm{E}[X|r_x]$ would increase less rapidly with further increases in r_x, for values of r_x above $\log 50$. This would result in a frequent failure to recognize how attractive the risky lottery truly is when X exceeds \$50, and hence less frequent acceptance of the risky lottery in such cases than the scale-invariant model would predict, as can be observed in Figure 2. (We leave for future work more detailed consideration of the extent to which our data may be better explained by a more subtle account of subjects' prior beliefs, or by a model of noisy coding that is not exactly logarithmic.)

Holt and Laury (2002) also obtain nearly perfect scale-invariant choice curves (see their Figure 1), when the amounts offered in hypothetical gambles are scaled up by a factor as large as 90 times those used in their small-stakes gambles. They find, however, that their subjects' apparent degree of risk aversion increases when the scale of the gambles is increased, in the case of gambles for real money (their Figure 2). It is unclear whether this difference from our results (which also involve real money) reflects a difference in the kind of gambles presented to their subjects, or the fact that their large gambles involved greater amounts of money than even our largest gambles (hundreds of dollars rather than mere tens of dollars).[21] Further studies would be desirable to clarify this.

3.2. *Comparison with random expected utility models*

As noted in the introduction, both the random variation in subjects' choices between simple gambles and existence of small-stakes risk aversion are often explained, in the experimental economics literature, by positing (i) "narrow bracketing" of the choice problem, so that the small amounts that can be gained in the experiment are not integrated with the subject's overall wealth (or overall lifetime budget constraint), (ii) significant concavity of the utility function that is used to value different possible monetary gains in the experiment, and (iii) a random term in the utility function, so that the expected utility assigned to a given probability distribution over possible gains is not always the same. We have offered an alternative model of both the randomness

20. Note however that this curve is also less well estimated than the others shown in the figure, as a number of our subjects were not presented with trials including values of C this large, so that the N_{obs} for this case is smaller, as indicated in Table 3 in the Supplementary Appendix.

21. Note that it is perfectly consistent with our model to suppose that diminishing marginal utility of wealth becomes an additional source of risk aversion in the case of large gambles, as we would expect to be the case. It should also be noted that our model implies scale-invariance under the assumption that the DM's prior should be the same on the trials with different values of C; this makes sense in the context of our experiment (where trials with different values of C are randomly interleaved), but less obviously so in the experiments of Holt and Laury.

TABLE 2
In-sample and out-of-sample measures of goodness of fit for three models: our logarithmic coding model and two additive random-utility models. The format is the same as in Table 1. (See text for further explanation.)

Model	$LL^{\text{calibration}}$	BIC	$LL^{\text{validation}}$	log K
		Pooled data		
Log coding	−2,812.7	5,642.9	−940.0	0.0
ARUM-Probit	−2,997.3	6,012.0	−1,001.8	246.3
ARUM-Logit	−2,973.4	5,964.2	−993.5	214.1
		Heterogeneous parameters		
Log coding	−1,853.4	3,932.6	−677.5	0.0
ARUM-Probit	−1,960.0	4,145.8	−763.2	192.3
ARUM-Logit	−1,903.7	4,033.1	−688.0	60.7

and the degree of apparent risk aversion in the choices of our subjects that we regard as more theoretically parsimonious, and in our view this theoretical parsimony should be a reason to prefer our interpretation, even if the competing views were equally consistent with the data from a single experiment such as this one. Nonetheless, it is interesting to ask whether our data could not be equally well explained by a more familiar model.

Table 2 compares the fit of our model with two variants of an additive random-utility model (ARUM). In the case of each of the ARUMs, the subject is assumed to choose the option for which $\mathrm{E}[u(Y)]+\epsilon$ is larger, where Y is the monetary amount gained from the experiment (a random quantity, in the case of a risky prospect), $u(Y)$ is a non-linear utility function for such gains (valued separately from the subject's other wealth), and ϵ is a random term (drawn at the time of choice) that is independent of the characteristics of the option, and also independent of the corresponding random term in the value assigned to the other option. In the ARUMs considered in the table, $u(Y)$ is assumed to be of the constant relative risk aversion (CRRA) form, $u(Y)=Y^{1-\gamma}/(1-\gamma)$, for some $\gamma \geq 0$. The random term ϵ is assumed either to be normally distributed (the ARUM-Probit model), or to have an extreme-value distribution (the ARUM-Logit model). Thus each of the ARUMs has two free parameters (the coefficient of relative risk aversion γ and the standard deviation of ϵ), like the logarithmic coding model.

As in the case of Table 1, we consider both in-sample and out-of-sample measures of model fit, where the calibration dataset and validation dataset are the same as in the earlier table. In each case, we find that our model based on logarithmic coding provides a better fit to the experimental data, both in-sample and out-of-sample. The alternative model which comes closest to being a competitor is the ARUM-logit model, when separate parameters are estimated for each subject. Yet even in this case, the implied Bayes factor $K > 10^{26}$. If one of the models considered must represent the correct statistical model of our data, then the evidence overwhelmingly favours the model based on logarithmic coding.

In the Supplementary Appendix, we report similar statistics for additional variants. Allowing $U(Y)$ to belong to the more general class of hyperbolic absolute risk aversion (HARA) utility functions does not improve the fit of the ARUMs, once the penalty for the additional free parameter is taken into account. We also consider an alternative model of stochastic choice proposed by Apesteguia and Ballester (2018), based on EUM with a random parameter in the utility function, and show that this formulation does not better explain our data, either. The ARUMs above can also be considered random variants of prospect theory, in which $u(Y)$ is the Kahneman–Tversky value function for gains, but we use the true probabilities of the two outcomes as weights rather than distorted weights of the kind posited by Kahneman and Tversky (1979). In the Supplementary Appendix, we show that allowing for a probability weight different from the true probability does not improve the fit of the random version of prospect theory, once the penalty for the additional free parameter is taken into account. Thus our model compares well with many of the most frequently used empirical specifications.[22]

4. FURTHER IMPLICATIONS OF THE THEORY

We have shown that it is possible to give a single unified explanation for the observed randomness in choices by subjects evaluating risky income prospects on the one hand, and the apparent risk aversion that they display on average on the other, as natural consequences of people's intuitions about the value of gambles being based on imprecise internal representations of the monetary amounts that are offered. Our theory explains the possibility of small-stakes risk aversion without implying any extraordinary degree of aversion to larger gambles in other contexts. Moreover, it can also explain the fact (demonstrated in our experiment) that the degree of risk aversion, as measured by the percentage by which the expected value of a random payoff must exceed the certain payoff in order for a subject to be indifferent between them, is relatively independent of the size of the stakes (as long as these remain small), contrary to what should be found if risk aversion were due to diminishing marginal utility.

We have argued in the introduction that we find this theory particularly compelling on account of its parsimony: the introduction of a new parameter (the parameter ν indicating the degree of imprecision of a subject's internal representation of monetary amounts) in order to account for the degree of randomness of a subject's choices also immediately implies their apparent degree of risk aversion, rather than this being controlled by a separate parameter. Moreover, the same basic theory (that intuitive judgments are made in accordance with an optimal decision rule, subject to their having to be based on noisy internal representations of the decision problem) has implications for a number of other types of decision problems, beyond the simple type of experiment considered in Sections 2 and 3. In particular, our model can also account for a number of other anomalous features of subjects' choices with regard to small gambles that are documented by Kahneman and Tversky (1979) and Tversky and Kahneman (1992).

The "reflection effect." Kahneman and Tversky report that if subjects must choose between a risky loss and a certain loss—with similar probabilities and monetary quantities as in the kind of problem considered above, but with the *signs* of the monetary payoffs reversed—risk *seeking* is observed more often than risk aversion (something they call the "reflection effect"). The coexistence of both risk-averse choices and risk-seeking choices by the same subject, depending on the nature of the small gambles that are offered, is a particular puzzle for the EUM account of risk attitudes, since a subject should be either risk averse or risk seeking (depending whether the subject's utility of wealth is concave or convex) regardless of the sign of the gambles offered.

The explanation of risk aversion for small gambles offered here instead naturally implies that the sign of the bias (i.e., of the apparent risk attitude) should switch if the signs of the monetary payoffs are switched. Consider instead the case of a choice between a risky gamble that offers a probability p of losing an amount X (but losing nothing otherwise), and the option of a certain loss of an amount C. If we assume that the quantities X and C are mentally represented according to the same logarithmic coding model as above, regardless of whether they represent gains or losses, then in the case of losses, the subject's expected wealth is maximized by a rule under

22. In the Supplementary Appendix, we also discuss the degree to which our data are consistent with random versions of the model proposed by Bordalo *et al.* (2012), in which risk attitudes result from differences in the salience of alternative outcomes. The consistency of such a model with our data depends greatly on the way in which the qualitative and deterministic model in their paper is made quantitative and stochastic.

which the risky lottery is chosen if and only if

$$p \cdot \mathrm{E}[X|r_x] < \mathrm{E}[C|r_c], \tag{4.1}$$

reversing the sign in (2.2).

The set of internal representations (r_x, r_c) for which this holds will be the complement of the set discussed earlier, so that the model predicts

$$\text{Prob}[\text{accept risky}|X,C] = \Phi\left(\frac{\beta^{-1}\log p^{-1} - \log X/C}{\sqrt{2}\nu}\right). \tag{4.2}$$

Indifference again will require $pX > C$, but this will now count as *risk-seeking* behaviour; when $pX = C$, the risky loss should be chosen more often than not.

This explanation for the "reflection effect" is not fundamentally different from those of Kahneman and Tversky (1979) or Bordalo *et al.* (2012), who attribute it to the fact that diminishing marginal sensitivity exists for losses as well as for gains. The additional insight offered by our model is its provision of a further account of the origin of diminishing marginal sensitivity, and in particular, the demonstration that it is consistent with a hypothesis that subjects' responses maximize the utility that subjects obtain (on average) from the use of their total monetary wealth, without any assumption of an intrinsic concern with gains or losses.

A framing effect. Kahneman and Tversky (1979) further show that subjects' preferences between a risky and a safe outcome can be flipped, depending whether the options are presented as involving gains or losses. In one of their problems, subjects are asked to imagine being given a substantial monetary amount $2M$, and then being presented with a choice between (a) winning an additional M with certainty, or (b) a gamble with a 50% chance of winning another $2M$ and a 50% chance of winning nothing. In a second problem, the initial amount was instead $4M$, and the subsequent choice was between (a) losing M with certainty, and (b) a gamble with a 50% chance of losing $2M$ and a 50% chance of losing nothing.

These two problems are equivalent, in the sense that in each case the subject chooses between (a) ending up with $3M$ more than their initial wealth with certainty, or (b) a gamble under which they have an equal chance of ending up with $2M$ or $4M$ more than their initial wealth. Nonetheless, a substantial majority of their subjects chose (a) in the first problem, while a substantial majority chose (b) in the second. This contradicts any theory (not just EUM) under which people should have a consistent preference ranking of probability distributions over final wealth levels.

Our theory easily explains this finding. If the initial gift is denoted G, and the monetary amounts G, X, and C defining the decision problem must each be independently represented in the fashion postulated above, then in the first problem, an expected wealth-maximizing decision rule will choose (b) if and only if

$$\mathrm{E}[G|r_g] + p \cdot \mathrm{E}[X|r_x] > \mathrm{E}[G|r_g] + \mathrm{E}[C|r_c],$$

which is equivalent to (2.2), while in the second problem it will choose (b) if and only if

$$\mathrm{E}[G|r_g] - p \cdot \mathrm{E}[X|r_x] > \mathrm{E}[G|r_g] - \mathrm{E}[C|r_c],$$

which is equivalent to (4.1). We then get different probabilities of choosing (b) in the two cases, given by equations (2.5) and (4.2), respectively.

Note that our theory assumes that the decision rule is in all cases the one that maximizes expected final wealth, so that only the sum of the initial gift and the additional gain or loss from the further option is assumed to matter to the decision maker; there is no intrinsic interest assumed in gains or losses relative to what one had in the past or what one expected to have. The relevance of the sequence of gains and losses by which one arrives at a given final wealth comes not from the decision maker's assumed objective, but from the need to mentally represent the quantities used in the description of the options, in the form in which they are presented, before integrating the separate pieces of information in further calculations. If this representation were possible with infinite precision (and subsequent operations could also be perfectly precise), then different ways of presenting information that imply the same possible final wealth levels would indeed be equivalent, and lead to the same choices. But when the precision with which each monetary amount can be represented is limited, mathematically equivalent problems are not processed in identical ways, and the resulting behaviour can be different as a result, despite its optimality in each case (conditional on the mental representation).

Loss aversion. In the discussion above, we assume that the unsigned magnitudes of monetary losses are encoded and decoded in exactly the same way as the magnitudes of monetary gains. While this leads to an especially parsimonious mathematical specification, the logic of our theory does not require us to assume this. In particular, the model should only be symmetric in the way that losses as opposed to gains are treated if the prior distribution from which potential losses are drawn is assumed to be the same as the prior distribution for potential gains; yet it is hardly obvious that this is true. If we assume that both gains and losses are log-normally distributed, but with different parameterizations of the two priors, then the theory developed above can be generalized (as shown in the Supplementary Appendix) to one in which the biases in the average subjective valuations of losses differ in size from those in the case of subjective valuations of gains. In particular, parameterizations of the model are possible in which it predicts that subjects should exhibit "loss aversion" of the kind also posited by prospect theory. For example, in order for subjects to be indifferent between accepting and rejecting a gamble that offers an equal chance of a gain of size G and a loss of size L, it may be necessary for the ratio G/L to be well above 1, as found by Tversky and Kahneman (1992).

As shown in the Supplementary Appendix, the parametric assumptions about the respective prior distributions for gains and losses required for this to be true are consistent with the assumptions needed to obtain the predictions regarding choices between risky gains and certain gains that we test above, as well as those needed to obtain the predictions regarding choices between risky losses and certain losses that we have just discussed. Whether the assumptions regarding priors that would be needed are also reasonable ones for settings such as the experiment reported by Tversky and Kahneman is a topic that deserves further elaboration in future studies.

Imprecise representation of probabilities. Tversky and Kahneman (1992) document other respects in which subjects make both risk-averse choices and risk-seeking choices with respect to small gambles, depending of the nature of the problem. For example, their subjects are more often risk-seeking when choosing between a small certain gain and a small probability of a considerably larger gain; but they are more often risk-averse when choosing between a modest certain loss and a small probability of a considerably larger loss. Prospect theory explains such cases by postulating that in the case of a risky prospect, the values assigned to the various possible gains or losses are weighted not in proportion to each outcome's probability of occurrence (as required by EUM), but rather using weights that represent a non-linear transformation of the true probabilities.[23]

Choices of this kind are consistent with our model, if the model is generalized to assume (in the case of simple gambles of the kind discussed above) that the probability p of the non-zero

23. Prelec (1998) and Gonzalez and Wu (1999) provide important further discussions of the properties that such a non-linear transformation should satisfy.

outcome also has an internal representation r_p that is probabilistic. In the analysis above, we have assumed for simplicity that the exact value of p is available as an input to the decision rule. This is not inconsistent with our general view of the internal representation of numerical information; for in the situation considered in our theoretical model (and in our experiment), there is a single value of p that is used in all trials (though X and C vary from trial to trial). Thus, if we assume that the prior for which the subject's decision rule has been optimized represents the distribution of possible decision situations *in this specific experiment*, the prior (in this type of experiment) would allow for only a single value of p—and the Bayesian posterior would similarly have this single value of p in its support, regardless of the noisy representation r_p.

Nonetheless, it is clearly of interest to consider the case in which the prior is non-degenerate (either because p varies from trial to trial, or because the decision maker has not had enough experience with a particular context for her decision rule to be adapted to the precise statistics of that context). Let us again assume for simplicity that under the prior, the quantities p, X, and C are distributed independently of one another; that the distribution of the representation r_p depends only on p (not on the values of X or C), and that the conditional distribution of r_p is independent of the realizations of r_x or r_c; and similarly for the other components of the internal representation $\mathbf{r}$. Then condition (2.2) for an optimal decision rule takes the more general form

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

or alternatively (given the model proposed above for the noisy coding of monetary amounts),

$$r_x - r_c > \beta^{-1}\rho, \quad \rho \equiv -\log \mathrm{E}[p|r_p], \tag{4.3}$$

generalizing (2.4). Here, ρ is a random variable (because it depends on the realization of r_p), conditional on the true probability p.

This generalization of the decision rule (2.2) results in an additional source of *randomness* in choice (namely, random variation in ρ); but in general, it also results in an additional source of *bias* (because ρ also differs from $\log p$ on average). As a simple example, suppose that the noise in the internal coding of X and C is negligible (ν is extremely small), but that the internal representation r_p is drawn from a distribution

$$r_p \sim N(z(p), \nu_p^2),$$

where $z(p) \equiv \log(p/1-p)$ are the log odds of the two outcomes, and ν_p is non-negligible. Suppose furthermore that the log odds are normally distributed under the prior,

$$z(p) \sim N(\mu_p, \sigma_p^2).$$

Then the posterior log odds, conditional on the representation r_p, are also normally distributed, with a mean $\bar{m}(r_p)$ that is a weighted average of r_p and the prior mean log odds, and a variance $\bar{\sigma}^2$ that is independent of r_p.

In this case, (4.3) requires that the value of $\log(X/C)$ required for indifference (acceptance of the gamble exactly half the time) will be equal to the median value of ρ, which (given that r_p has a symmetric distribution) is the value of ρ when r_p is equal to the true log odds. Alternatively, the value of the ratio C/X required for indifference is given by a function $w(p)$, where

$$w(p) \equiv \mathrm{E}[F(z(p), \epsilon)], \qquad F(z,\epsilon) \equiv \frac{\exp[\bar{m}(z)+\epsilon]}{1+\exp[\bar{m}(z)+\epsilon]},$$

and ϵ is a random variable distributed $N(0, \bar{\sigma}^2)$. This function plays a role similar to the probability weighting function of Kahneman and Tversky (1979). And as long as the variance $\bar{\sigma}^2$ is not too great, the model implies that $w(p)$ will have the inverse-S shape assumed by Kahneman and Tversky.

Both Tversky and Kahneman (1992) and Gonzalez and Wu (1999) argue that this shape for the probability weighting function can be understood as another example of diminishing marginal sensitivity—in this case, diminishing with distance from either of the two extreme probabilities 0 and 1. Our model shows more precisely how the same principle (noisy coding combined with Bayesian inference from the noisy internal representation) naturally gives rise to diminishing marginal sensitivity to larger monetary gains or losses and to effectively non-linear probability weighting.

In particular, if we fix the ratio ν_p/σ_p but make both σ_p and ν_p small, then in the limit as $\sigma_p, \nu_p \to 0$, we obtain an analytical solution of the form

$$w(p) = \frac{\alpha p^\beta}{(1-p)^\beta + \alpha p^\beta}, \tag{4.4}$$

for certain coefficients $\alpha > 0, 0 < \beta < 1$, which depend on the ratio ν_p/σ_p and the prior mean log odds μ_p. (See the Supplementary Appendix for details.) This function has the inverse-S shape assumed by Kahneman and Tversky; indeed, the two-parameter family of weighting functions (4.4) has often been assumed in econometric implementations of prospect theory.[24]

In this case, the model predicts risk-seeking in the case of gains for low values of p, but risk-aversion for larger values of p, and risk-aversion in the case of losses for low values of p, but risk-seeking for larger values of p, all as found by Kahneman and Tversky (1979). The model is also consistent with the occurrence of Allais-type paradoxes, as discussed further in the Supplementary Appendix.

5. EVIDENCE OF A CONNECTION BETWEEN RISK AVERSION AND COGNITIVE IMPRECISION

Thus far, we have argued that our theory offers a unified explanation of a wide range of phenomena that are not only well-documented, but that, in previous accounts, have required a number of seemingly unrelated features to be added to a descriptive model of choice behaviour. However, we have discussed only phenomena that are also predicted by a variety of already existing models of choice under risk. A further virtue of our theory is that it also leads to a variety of predictions that would not be predicted by other accounts, that do not tie risk attitudes with respect to small gambles to cognitive imprecision.

Implications of heterogeneity in cognitive precision. We have shown above (Tables 1 and 2) that our experimental data are better fit by allowing the parameters of the scale-invariant psychometric function to vary across subjects. There is an obvious reason, in the context of our theory, for the choice curves of different subjects to be different: there is no reason to suppose that the parameter ν, indicating the degree of imprecision of internal representations of monetary amounts, is the same for all people. However, our theory implies that increasing ν should *both* increase the randomness of the subject's choices *and* imply greater apparent risk-aversion, as measured by the value of X/C required for indifference. To the extent that variation in this single factor is the main reason for differences in the behaviour observed for different subjects, we should expect to see a positive correlation between randomness of choice and degree of risk aversion.

24. Stott (2006) calls this family of weighting functions the "Goldstein-Einhorn" specification, and reviews the various authors (including some prior to the work of Kahneman and Tversky) who have proposed versions of it.

We can test this prediction by further examining the variation in behaviour across subjects. We focus here on the scale-invariant model, which as argued above is best supported by our data, and estimate a scale-invariant choice curve for each of the 19 subjects other than subject 9. (We omit subject 9 from the discussion in this section, since as shown in Figure 4, this subject's data are not well-described by a scale-invariant model.) Our theoretical model implies not only that a scale-invariant curve (3.2) should describe each subject's data, but also that the coefficients for each subject should satisfy the inequalities

$$\gamma \geq 0, \qquad -\frac{\delta}{\gamma} \geq \log p^{-1}, \tag{5.1}$$

which are required in order for there to exist values of σ^2 and ν^2 consistent with those coefficients. Hence in estimating the subject-specific models, we impose the further restriction that the value of γ be non-negative.[25]

In comparing the choice curves of the different subjects, it is useful to parameterize them not by γ and δ, but instead by the values of γ and

$$\pi \equiv e^{\delta/\gamma}.$$

In terms of this alternative parameterization, the theoretical constraints (5.1) can be written:

$$\gamma \geq 0, \qquad \pi \leq p. \tag{5.2}$$

Note that when $\delta/\gamma \leq 0$ (as required by our theoretical model), $0 \leq \pi \leq 1$, and π has the interpretation of a "risk-neutral probability": the subject's indifference point is the same as that of a risk-neutral, optimizing decision maker who believes that the probability of the non-zero lottery payoff is π (rather than the true probability p). The prediction that $\pi \leq p$ is another way of saying that our theoretical model predicts apparent risk aversion; and the degree to which a subject's estimated π is less than p provides a simple measure of the degree of apparent risk aversion.

Figure 5 shows the estimated values of γ and π for each of the 19 subjects for whom it is not grossly inaccurate to estimate a scale-invariant choice curve (now imposing the theoretical constraint that $\gamma \geq 0$). For each subject, an open circle indicates the parameter values that maximize the likelihood of the data for that subject (the maximum likelihood (ML) estimate), and the surrounding shaded region indicates the set of parameter values for which the log likelihood of the data is no more than 2 points lower than at the maximum. Thus the shaded region indicates a Bayesian credible region for the parameter estimates, under the assumption of a uniform prior for those parameters. The boundary of each maximum-posterior density credible region is chosen according to a criterion which, in the case of a Gaussian posterior for a single variable, would report the interval corresponding to the mean estimate plus or minus two standard deviations.

The largest value of π that would be consistent with prediction (5.2) is indicated by the horizontal dotted line in Figure 5; we see that for all but one of the 19 subjects, the credible region includes points consistent with this prediction. Thus, the individual choice curves of 18 out of our 20 subjects are reasonably consistent with both the model prediction of scale invariance and with the coefficient constraints (5.2).

25. For all but one subject, the estimated value of γ would be positive, even without imposing the restriction, and even for that one, the value of γ is only slightly negative; we presume that this estimate reflects sampling error, and treat this subject as having a γ of zero.

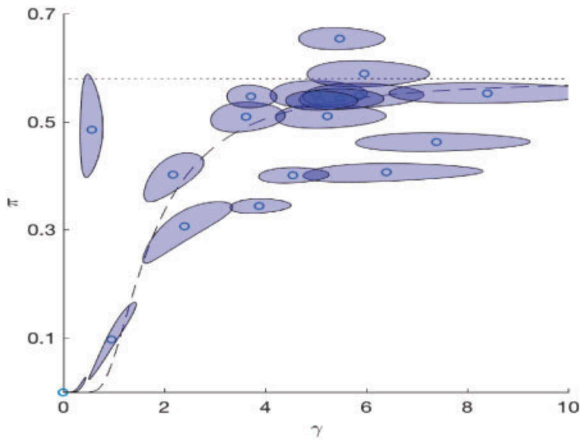

FIGURE 5

Heterogeneity in subjects' choice curves. Each shaded region indicates the credible region for an individual subject's parameters γ and π, with an open circle at that subject's maximum-likelihood values. The dashed line shows the theoretical relationship between γ and π that should exist if all subjects share a common prior, under which $\sigma = 0.35$.

If we further assume a common log-normal prior (2.3) for all subjects, but allow the precision of mental coding of monetary amounts (*i.e.* the parameter ν^2) to vary across subjects, then the values of γ and π estimated for each subject are predicted by the model to be linked by a relationship of the form

$$\pi = p^{1+(2\sigma^2\gamma^2)^{-1}}, \tag{5.3}$$

where σ^2 is a parameter common to all subjects. This is an upward-sloping relationship, of the kind illustrated by the dashed curve in Figure 5, which graphs equation (5.3) in the case that $\sigma = 0.35$. Here, ν^2 is decreasing (the precision of mental coding is increasing) as one moves up and to the right along the dashed curve.

If we estimate a choice curve for each subject without imposing the restriction of a common σ^2, the estimated coefficients do not all line up perfectly on a curve consistent with (5.3); nonetheless, there is a strong positive correlation between the ML estimates of γ and π for the various subjects, as can be seen in Figure 5. That is, the degree of apparent risk aversion (measured by the degree to which π is less than p) is generally greater for those subjects whose choices are less sensitive to variation in X/C (measured by the size of γ). The fact that these two features of behaviour go hand in hand is consistent with our theory. Models such as EUM or prospect theory, extended to allow for stochastic choice as in Section 3.2, instead provide no reason to expect such a relationship, since in these theories the degree of randomness of choice and the degree of risk aversion are determined by independent parameters.

The predictions of our theory will be more flexible—and arguably more consistent with our experimental data—if we allow different subjects to have different priors (more specifically, different values for σ^2). This might reflect imperfect learning of the prior (not surprising given the subjects' finite experience in this environment), in ways that differ across subjects.

Correlation with imprecision in number processing more generally. The positive correlation shown in Figure 5 is also found in a replication of our study by Garcia *et al.* (2018). They show furthermore that both the randomness and the apparent risk aversion in choice under risk can be predicted by the degree of randomness of the subject' responses in an independent numerosity comparison task of the kind discussed in Section 1.1—that is, by the subject-specific value of ν estimated using equation (1.2). This not only supports the hypothesis that small-stakes risk aversion results from noisy coding of monetary amounts, but suggests that the cognitive imprecision involved may be related to imprecision in the internal representation of numerical magnitudes more generally.

Schley and Peters (2014) also report greater apparent risk aversion in subjects who represent numbers less precisely in other types of tasks. These authors also offer an explanation for apparent risk aversion that is based on the idea that the perception of the numerical magnitudes of prospective monetary payoffs is biased, and more specifically that perceived magnitudes are an increasing, strictly concave function of the magnitudes. Like us, they base their proposal on limitations on people's general ability to accurate represent numbers mentally, rather than on the true utility obtained from having more money (as in the EUM explanation of risk aversion) or a theory of distorted valuations that is specific to the domain of value-based decision making (as with prospect theory). In support of this proposal, they show that subjects who less accurately represent numbers for other purposes also exhibit greater apparent diminishing marginal utility of income and greater apparent risk-aversion in choices between risky gambles.

However, their discussion assumes that less capacity for symbolic number mapping results in a deterministic distortion of perceived numerical magnitudes (a true quantity X is always perceived as exactly $\hat{X} = AX^{\beta}$), rather than in a more random mental representation as in our theory. This means that they do not explore the connection between the randomness of subjects' choices and apparent risk aversion, as Garcia *et al.* and we do; and their theory provides no explanation for why people should value lotteries according to the average value of the perceived payoffs $\hat{X}_i$ instead of, say, according to the average value of $\hat{X}_i^{1/\beta}$—a criterion that would reliably maximize expected wealth, taking into account the perceptual distortion.

Effects of varying cognitive load. Thus far, we have discussed implications of our model, taking the precision of coding (parameterized by ν) to be fixed. But the model also makes predictions about the effects of varying ν, which might be subject to predictable variation for a variety of reasons. For example, one might well suppose that increased time pressure, distraction, or cognitive load should reduce the cognitive resources used to represent the monetary magnitudes that define a particular decision problem, and that this should correspond, in our model, to an increase in ν. According to our model, this should result in both decreased sensitivity of a subject's decisions to variations in the risky payoff X that is offered (*i.e.* a lower value of γ) and increased apparent risk aversion (a value of π that is lower relative to p). This is an example of a prediction of our theory that is not made by theories like prospect theory, that attribute departures from the predictions of EUM to (presumably stable) distortions in the way that subjects evaluate monetary gains or probabilities.

In fact, a number of authors have found that increasing cognitive load (*e.g.* by requiring subjects to concurrently maintain a list of random letters or numbers in memory) causes subjects to make more risk-averse choices (Whitney *et al.* 2008; Benjamin *et al.* 2013; Deck and Jahedi 2015;

Gerhardt *et al.* 2016).[26] This is often interpreted as support for a "dual systems" view of decision making, in which increased cognitive load makes it harder for subjects to employ a deliberative system that would be called upon under other circumstances, so that emotional reactions or simpler heuristics are relied upon instead. Our theory provides an alternative interpretation, in which the same cognitive mechanism might be employed in both cases, but it relies upon an imprecise analog representation with a degree of precision that depends on the number of other claims on working memory. The fact that our subjects display a range of degrees of apparent risk aversion, as well as a range of degrees of randomness in their choices, as shown in Figure 5—rather than simply two clusters corresponding to the users of two very different mental systems—is more easily explained under the theory that we propose.

Effects of varying the range of possible monetary payoffs. Another reason for the parameter ν to possibly vary across experimental settings is change in the ranges of monetary

payoffs X and C that need to be represented. While we have found that we can account fairly well for the imprecision in our subjects' evaluations of the gambles presented to them under the hypothesis that the coding noise parameter ν is the same on all trials, it remains possible that the value of ν adapts depending on the prior distribution from which X and C are expected to be drawn. Frydman and Jin (2019) report an experiment similar to ours, in which however the parameter σ of the log-normal distributions from which the monetary amounts X and C are drawn is different in two different blocks of trials. They observe probabilistic dependence of choice on the payoffs offered on a given trial (as in our Figures 2 and 3), but find that the choice curves are steeper (when plotted as a function of $\log(X/C)$) in the data from the sessions in which σ was smaller. They interpret this result as showing that the precision of internal representations of the monetary payoffs depends on the range of values that the decision maker expects to encounter in a given context, in a way that allows the brain to make optimal use of a finite capacity for discriminating between different amounts ("efficient coding").

Our theory is easily extended to predict such an effect. Suppose that the limit on the precision of numerical representations comes not from some "hard-wired" system of number representation (so that the degree of overlap between the distributions of internal representations of the quantities \$28 and \$30, say, is independent of the prior distribution from which these quantities may have been drawn), but from a limit on the information-processing capacity of the system used to represent analog magnitudes. As a concrete example of such a theory, suppose that the representation r_x of any quantity X is given by an independent draw from a distribution $N(m,\omega^2)$, where $m=\phi(X)$ is a non-linear transformation of the true magnitude. If we assume that production of representations with this degree of fidelity requires that the variance of the distribution of values for m be less than some finite upper bound Ω^2, then the system has a finite information-theoretic "channel capacity" determined by the ratio Ω/ω — essentially, a measure of the effective number of distinct cases that can be accurately discriminated. (See Cover and Thomas, 2006, chap. 9.)

Let us suppose that ω and Ω are given parameters (reflecting the limited acuity of a person's semantic representations of numerical magnitudes), but that the transformation $\phi(X)$ is optimized for a given environment. More specifically, suppose that the transformation is of the form $\phi(X)=\xi+\psi\log X$, with parameters ξ,ψ that are optimized for a given environment (that is, for a given prior distribution for X). We show in the Supplementary Appendix that if the prior is of the

26. Olschewski *et al.* 2018 instead find that increased cognitive load increases the randomness of choice, but only increases risk aversion by a small (statistically insignificant) amount. Their measure of risk aversion, however, is based on an "ARUM-probit" model that, as shown in Section 3.2, is not consistent with the predictions of our model, and does not fit our own experimental data as well as the logarithmic coding model.

form $X \sim N(\mu, \sigma^2)$, then the optimal encoding rule leads to a noisy internal representation of the kind assumed in section 2, but with a value of ν that varies endogenously with the prior. In fact, $\nu = (\omega/\Omega)\sigma$, so that ν is predicted to grow in proportion to the value of σ. Combining this result with prediction (2.5) for the probability of acceptance of the risky bet, we see that the model predicts the effect observed by Frydman and Jin (2019), in which a lower value of σ results in a steeper choice curve.[27]

An important implication of both the results of Frydman and Jin, and the studies suggesting that manipulations of cognitive load can affect apparent risk aversion, is that an aspect of choice behaviour commonly attributed to preferences—which are usually assumed to remain invariant across choice situations—may actually be malleable: subject to variation from one situation to another, and potentially subject to systematic influence by the designer of a choice environment.[28] The extent to which this is true should be an important topic for further study.

The theory proposed here (and the experimental results that we report) neither requires that this be the case, nor precludes it; our main results simply assume a particular model of (and degree of precision of) noisy coding of numerical information about the choice options, without asking where it came from, and assume a decision rule that is optimal for a particular prior distribution over choice situations, without asking how the optimal rule (or the prior relative to which it is optimal) is learned. But exactly because we seek to ground risk taking in more general features of cognition rather than appealing to preferences, it is natural to consider how these features of a given person's cognitive functioning may be shaped by their experience.[29] Extension of the theory to incorporate learning dynamics and adaptation of internal representations to the statistics of a given environment should be an important next step in the research agenda.

Acknowledgments. An earlier version of this work, under the title "Cognitive Limitations and the Perception of Risk," was presented as the 2015 AFA Lecture at the annual meeting of the American Finance Association. We thank Colin Camerer, Tom Cunningham, Daniel Friedman, Cary Frydman, Drew Fudenberg, Xavier Gabaix, Nicola Gennaioli, Frank Heinemann, Lawrence Jin, Arkady Konovalov, David Laibson, Ifat Levy, Rosemarie Nagel, Charlie Plott, Rafael Polania, Antonio Rangel, Christian Ruff, Andrei Shleifer, Hrvoje Stojic, Chris Summerfield, Shyam Sunder, Peter Wakker, Ryan Webb, and four anonymous referees for helpful comments, and the National Science Foundation for research support.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ANOBILE, G., MARCO CICCHINI, G. and BURR, D. C. (2012), "Linear Mapping of Numbers onto Space Requires Attention," *Cognition* **122**, 454–459.

APESTEGUIA, J. and BALLESTER, M. A. (2018), "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences," *Journal of Political Economy* **126**, 74–106.

ARROW, K. J. (1971), *Essays in the Theory of Risk-Bearing* (Chicago: Markham Publishing Co.).

BANKS, W. P., FUJII, M. and KAYRA-STEWART, F. (1976), "Semantic Congruity Effects in Comparative Judgments of Magnitudes of Digits," *Journal of Experimental Psychology: Human Perception and Performance* **2**, 435–447.

BAYRAK, O. K. and HEY, J. D. (2019), "Understanding Preference Imprecision" (Working Paper).

27. Frydman and Jin (2019) also propose a model of efficient coding that can explain their finding. However, their formulation has other quantitative predictions that are at odds with our experimental data; for example, their model implies that choice curves should not be scale-invariant. The efficient coding model proposed here is instead consistent with both our findings and theirs.

28. See Payzan-LeNestour and Woodford (2020) for further discussion. "Evolutionary" theories of the origin of preferences, such as those of Robson (2001), Rayo and Becker (2007), and Netzer (2009), suggest that this could be possible as well, though it is often unclear over what time scale the "evolution" is thought to occur in such theories.

29. For an example of a model of a dynamic process through which a noisy coding scheme might adapt to a changing environment, see Robson and Whitehead (2019).

BENJAMIN, D. J., Brown, S. A. and SHAPIRO, J. M. (2013), "Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Journal of the European Economics Association* **11**, 1231–1255.

BLOCK, H. D. and MARSCHAK, J. (1960), "Random Orderings and Stochastic Theories of Response", in Olkin, I. *et al.,* (Eds), *Contributions to Probability and Statistics* (Stanford: Stanford University Press).

BORDALO, P., GENNAIOLI, N. and SHLEIFER, A. (2012), "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics* **127**, 1243–1285.

BRANNON, E. M. (2006), "The Representation of Numerical Magnitude," *Current Opinion in Neurobiology* **16**, 222–229.

BUCKLEY, P. B. and GILLMAN, C. B. (1974), "Comparisons of Digits and Dot Patterns," *Journal of Experimental Psychology* **103**, 1131–1136.

BURNHAM, K. P. and ANDERSON, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach,* 2d. ed., New York: Springer.

BUTLER, D., ISONI, A. and LOOMES, G. (2012), "Testing the 'Standard' Model of Stochastic Choice Under Risk," *Journal of Risk and Uncertainty* **45**, 191–213.

BUTLER, D. and LOOMES, G. (1988), "Decision Difficulty and Imprecise Preferences," *Acta Psychologica* **68**, 183–196.

BUTLER, D. and LOOMES, G. (2007), "Imprecision as an Account of the Preference Reversal Phenomenon," *American Economic Review* **97**, 277–297.

CANTLON, J. F. and BRANNON, E. M. "Shared System for Ordering Small and Large Numbers in Monkeys and Humans," *Psychological Science* **17**, 401–406.

CORDES, S., GELMAN, R., GALLISTEL, C. R. et al. (2001), "Variability Signatures Distinguish Verbal from Non-Verbal Counting for Both Large and Small Numbers," *Psychonomic Bulletin and Review* **8**, 698–707.

COX, J. C., SADIRAJ, V., VOGT, B. and DASGUPTA, U. (2013), "Is There a Plausible Theory for Decision Under Risk? A Dual Calibration Critique," *Economic Theory* **54**, 305–333.

DECK, C. and JAHEDI, S. (2015), "The Effect of Cognitive Load on Economic Decision Making: A Survey and New Experiments," *European Economic Review* **78**, 97–119.

DEHAENE, S. (1992), "Varieties of Numerical Abilities," *Cognition* **44**, 1–42.

DEHAENE, S. (2008), "Symbols and Quantities in Parietal Cortex: Elements of a Mathematical Theory of Number Representation and Manipulation", in Haggard, P., Rossetti, Y. and Kawato, M. (eds) *Sensorimotor Foundations of Higher Cognition* (Oxford: Oxford University Press).

DEHAENE (2011), *The Number Sense.* Revised and updated edition. (Oxford: Oxford University Press).

DEHAENE, S. and COHEN, L. (1991), "Two Mental Calculation Systems: A Case Study of Severe Acalculia with Preserved Approximation," *Neuropsychologia* **29**, 1045–1074.

DEHAENE, S., DUPOUX, E. and MEHLER, J. (1990), "Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison," *J. Experimental Psychology: Human Perception and Performance* **16**, 626–641.

DEHAENE, S. and FREDERICO MARQUES, J. (2002), "Cognitive Euroscience: Scalar Variability in Price Estimation and the Cognitive Consequences of Switching to the Euro," *Quarterly Journal of Experimental Psychology A* **55**, 705–731.

DEHAENE, S., PIAZZA, M., PINEL, P. et al., (2003), "Three Parietal Circuits for Number Processing," *Cognitive Neuropsychology* **20**, 487–506.

DRUGOWITSCH, J., WYART, V., DEVAUCHELLE, A. D. et al., (2016), "Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality," *Neuron* **92**, 1398–1411.

FECHNER, G. T. (1860), *Elements of Psychophysics* (Leipzig: Breitkopf under Bartel). [English translation published by Holt, Rinehart and Winston, 1966.]

FRIEDMAN, D., MARK ISAAC, R., JAMES, D. et al., (2014), *Risky Curves: On the Empirical Failure of Expected Utility* (London: Routledge).

FRYDMAN, C. and JIN, L. J. (2019), "Efficient Coding and Risky Choice" (Working Paper, University of Southern California).

FUDENBERG, D. and LEVINE, D. K. (2006), "A Dual-Self Model of Impulse Control," *American Economic Review* **96**, 1449–1476.

FUDENBERG, D. and LEVINE, D. K. (2011), "Risk, Delay, and Self-Control Costs," *American Economic Journal: Microeconomics* **3**, 34–68.

GARCIA, M., GRÜSCHOW, M., POLANIA, R. et al., (2018), "Predicting Risk Attitudes from the Precision of Mental Number Representation" *Poster Presented at the 16th Annual Meeting of the Society for Neuroeconomics, Philadelphia, PA, USA, October 5—7, 2018.*

GERHARDT, H., BIELE, G. P., HEEKEREN, H. R., et al., (2016), "Cognitive Load Increases Risk Aversion" (SFB 649 Discussion Paper no. 2016–011, Humboldt University Berlin).

GONZALEZ, R. and WU, G. (1999), "On the Shape of the Probability Weighting Function," *Cognitive Psychology* **38**, 129–166.

HEY, J. D. (1995), "Experimental Investigations of Errors in Decision Making under Risk," *European Economic Review* **39**, 633–640.

HEY, J. D. (2001), "Does Repetition Improve Consistency?" *Experimental Economics* **4**, 5–54.

HEY, J. D. and ORME, C. (1994), "Investigating Parsimonious Generalizations of Expected Utility Theory using Experimental Data," *Econometrica* **62**, 1291–1329.

HOLT, C. A. and LAURY, S. K. (2002), "Risk Aversion and Incentive Effects," *American Economic Review* **92**, 1644–1655.

INDOW, T. and IDA, M. (1977), "Scaling of Dot Numerosity," *Perception and Psychophysics* **22**, 265–276.

IZARD, V. and DEHAENE, S. (2008), “Calibrating the Mental Number Line,” *Cognition* **106**, 1221–1247.
JEVONS, W. S. (1871), “The Power of Numerical Discrimination,” *Nature* **3**, 281–282.
KAHNEMAN, D. and TVERSKY, A. (1979), “Prospect Theory: An Analysis of Decision Under Risk,” *Econometrica* **47**, 263–291.
KAUFMAN, E. L., LORD, M. W., REESE, T. W., et al., (1949), “The Discrimination of Visual Number,” *American Journal of Psychology* **62**, 498–525.
KHAW, M. W., LI, Z. and WOODFORD, M. (2020), “Replication Data for ‘Cognitive Imprecision and Small-Stakes Risk Aversion’,” Inter-University Consortium for Political and Social Research, Ann Arbor, MI, ICPSR120130V1. https://doi.org/10.3886/E120130V1.
KOSZËGI, B. and RABIN, M. (2008), “Revealed Mistakes and Revealed Preferences,” in Caplin, A. and Schotter, A. (Eds) *The Foundations of Positive and Normative Economics* (Oxford: Oxford University Press).
KRAMER, P., DE BONO, M. G. and ZORZI, M. (2011), “Numerosity Estimation in Visual Stimuli in the Absence of Luminance-Based Cues,” *PLoS One* **6**, e17378.
KRUEGER, L. E. (1972), “Perceived Numerosity,” *Perception and Psychophysics* **11**, 5–9.
KRUEGER, L. E. (1984), “Perceived Numerosity: A Comparison of Magnitude Production, Magnitude Estimation, and Discrimination Judgments,” *Perception and Psychophysics* **35**, 536–542.
LUYCKX, F., NILI, H., SPITZER, B. et al., (2018), “Neural Structure Mapping in Human Probabilistic Reward Learning,” *bioRXiv* preprint (doi: http://dx.doi.org/10.1101/366757), 10 July 2018.
MOSTELLER, F. and NOGEE, P. (1951), “An Experimental Measurement of Utility,” *Journal of Political Economy* **59**, 371–404.
MOYER, R. S. and LANDAUER, T. K. (1967), “Time Required for Judgements of Numerical Inequality,” *Nature* **215**, 1519–1520.
NETZER, N. (2009), “Evolution of Time Preferences and Attitudes Toward Risk,” *American Economic Review* **99**, 937–955.
NIEDER, A. (2013), “Coding of Abstract Quantity by ‘Number Neurons’ in the Primate Brain,” *Journal of Comparative Physiology A* **199**, 1–16.
NIEDER, A. and DEHAENE, S. (2009), “Representation of Number in the Brain,” *Annual Review of Neuroscience* **32**, 185–208.
NIEDER, A. and MERTEN, K. (2007), “A Labeled-Line Code for Small and Large Numerosities in the Monkey Prefrontal Cortex,” *Journal of Neuroscience* **27**, 5986–5993.
OLSCHEWSKI, S., RIESKAMP, J. and SCHEIBEHENNE, B. (2018), “Taxing Cognitive Capacities Reduces Choice Consistency Rather than Preference: A Model-Based Test,” *Journal of Experimental Psychology: General* **147**, 462–484.
PAYZAN-LENESTOUR, E. and WOODFORD, M. (2020), “Outlier Blindness: Efficient Coding Generates an Inability to Represent Extreme Values” (Working Paper).
PETZSCHNER, F. H., GLASAUER, S. and STEPHAN, K. E. (2015), “A Bayesian Perspective on Magnitude Estimation,” *Trends in Cognitive Sciences* **19**, 285–293.
PIAZZA, M., IZARD, V., PINEL, P. et al., (2004), “Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus,” *Neuron* **44**, 547–555.
PRELEC, D. (1998), “The Probability Weighting Function,” *Econometrica* **66**, 497–527.
RABIN, M. (2000), “Risk Aversion and Expected-Utility Theory: A Calibration Theorem,” *Econometrica* **68**, 1281–1292.
RABIN, M. and THALER, R. H. (2001), “Anomalies: Risk Aversion,” *Journal of Economic Perspectives* **15**, 219–232.
RAYO, L. and BECKER, G. S. (2007), “Evolutionary Efficiency and Happiness,” *Journal of Political Economy* **115**, 302–337.
ROBSON, A. (2001), “The Biological Basis of Economic Behavior,” *Journal of Economic Literature* **39**, 11–33.
ROBSON, A. J. and WHITEHEAD, L. A. (2019), “Adaptive Cardinal Utility” (Working Paper, Simon Fraser University).
ROSS, J. (2003), “Visual Discrimination of Number without Counting,” *Perception* **32**, 867–870.
SCHLEY, D. R. and PETERS, E. (2014), “Assessing ‘Economic Value’: Symbolic-Number Mappings Predict Risky and Riskless Valuations,” *Psychological Science* **25**, 753–761.
SPITZER, B., WASCHKE, L. and SUMMERFIELD, C. (2017), “Selective Overweighting of Larger Magnitudes During Noisy Numerical Comparison,” *Nature Human Behavior* **1**, 0145.
STOCKER, A. A. and SIMONCELLI, E. P. (2006), “Noise Characteristics and Prior Expectations in Human Visual Speed Perception,” *Nature Neuroscience* **9**, 578–585.
STOTT, H. P. (2006), “Cumulative Prospect Theory’s Functional Menagerie,” *Journal of Risk and Uncertainty* **32**, 101–130.
TEICHMANN, A. L., GROOTSWAGERS, T., CARLSON, T. et al. (2018), “Decoding Digits and Dice with Magnetoencephalography: Evidence for a Shared Representation of Magnitude,” *bioRXiv* preprint (doi: http://dx.doi.org/10.1101/249342) 23 January 2018.
THURSTONE, L. L. (1927), “A Law of Comparative Judgment,” *Psychological Review* **34**, 273–286.
TVERSKY, A. and Kahneman, D. (1992), “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty* **5**, 297–323.
VAN OEFFELEN, M. P. and VOS, P. G. (1982), “A Probabilistic Model for the Discrimination of Visual Number,” *Perception and Psychophysics* **32**, 163–170.
WEI, X.-X. and STOCKER, A. A. (2015), “A Bayesian Observer Model Constrained by Efficient Coding Can Explain ‘Anti-Bayesian’ Percepts,” *Nature Neuroscience* **18**, 1509–1517.

WEI, X.-X. and STOCKER, A. A. (2017), “Lawful Relation Between Perceptual Bias and Discriminability,” *Proceedings of the National Academy of Sciences United States of America* **114**, 10244–10249.

WHALEN, J., GALLISTEL, C. R. and GELMAN, R. (1999), “Non-Verbal Counting in Humans: The Psychophysics of Number Representation,” *Psychological Science* **10**, 130–137.

WHITNEY, P., RINEHART, C. A. and HINSON, J. M. (2008), “Framing Effects Under Cognitive Load,” *Psychonomic Bulletin and Review* **15**, 1179–1184.

WOODFORD, M. (2020), “Modeling Imprecision in Perception, Valuation and Choice,” *Annual Review of Economics,* **12**, 579—601.