Submitted to *Management Science*
manuscript MS-BDE-23-00265.R3

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

Decisions under Uncertainty as Bayesian Inference on Choice Options

Ferdinand M. Vieider
RISL$\alpha\beta$, Department of Economics, Ghent University
RISL$\alpha\beta$ Africa, AIRESS, University Mohammed VI Polytechnic

ferdinand.vieider@ugent.be

Standard models of decision-making under risk and uncertainty are deterministic. Inconsistencies in choices are accommodated by separate error models. The combination of decision model and error model, however, is arbitrary. Here, I derive a model of decision-making under uncertainty in which choice options are mentally encoded by noisy signals, which are optimally decoded by Bayesian combination with pre-existing information. The model predicts diminishing sensitivity towards both likelihoods and rewards, thus providing cognitive micro-foundations for the patterns documented in the prospect theory literature. The model is, however, inherently stochastic, so that choices and noise are determined by the same underlying parameters. This results in several novel predictions, which I test on one existing dataset and in two new experiments.

Key words: risk taking; noisy cognition; prospect theory;

> As far as the laws of mathematics refer to reality, they are not certain, as far as they are certain, they do not refer to reality.
>
> Albert Einstein (1922), p. 28

1. Motivation

Take a wager paying a prize x if event e occurs, and $y < x$ under the complementary event $\tilde{e}$. Assume that a decision maker has to choose between this wager and a sure amount of money, c, where $x > c > y$. Traditionally, decision makers have been assumed to have deterministic utilities over outcomes, as well as being able to form subjective beliefs over the occurrence of event e (Savage 1954). Choices of this type could then be used to recover the underlying belief and preference parameters of the decision maker. Empirical observations, however, have pointed to choices that are often inconsistent, even when the same choice options are repeated within relatively short time delays (Mosteller and Nogee 1951, Tversky 1969, Agranov and Ortoleva 2017).

The dominant response to such challenges in work aimed at identifying preferences from choice data has been to append an additive error term to the deterministic choice model (Hey and Orme 1994, Bruhin et al. 2010). The combination of error model and deterministic choice model, however, is arbitrary, in the sense that it provides many degrees of freedom to the econometric modeller. Buschena and Zilberman (2000) showed that inferences about the best-fitting choice model depend on the chosen error structure; and, vice versa, that conclusions about the most suitable error structure depend on the choice model adopted—a phenomenon that they describe as a 'path dependency problem for model selection' (p. 69). Conclusions reached based on any given combination of decision model and stochastic choice structure could thus be driven by arbitrary assumptions about the error structure (Alós-Ferrer et al. 2021).

I propose a model in which decision makers receive imperfect, noisy signals about the choice quantities. They then combine these noisy signals with mental priors describing the likelihood of different values in the environment to make inferences about the true choice options triggering the signals. The model is inherently stochastic, so that the decision model and stochastic choice model emerge organically from one and the same setup.[1] The model predicts decreasing sensitivity to rewards and probabilities much as documented in prospect theory (*PT*; Kahneman and Tversky 1979, Tversky and Kahneman 1992, Wakker 2010), thus establishing micro-foundations for observed choice behaviour. At the same time, however, the model departs markedly from PT in other respects. The same parameters determine discriminability between choice options *and* decision noise. The model thus predicts that the 'white noise' assumption underlying typical PT implementations is not warranted. This is problematic inasmuch as systematic correlations between errors and decision parameters may bias the inferences about preferences drawn from such models.[2]

I present several empirical tests to illustrate the model predictions. The first test examines novel predictions about patterns we should observe when estimating PT preference functionals on data that were truly generated by the Bayesian Inference Model (*BIM*) presented here. The first prediction concerns correlations between the decision parameters and noise. Perhaps surprisingly, such correlations have not been systematically examined to date. Correlations between noise and decision parameters may violate the 'PT plus white noise' assumption, and may affect the very inferences about preference parameters that are the main interest of such empirical estimations. The BIM, on the other hand, predicts exactly this sort of correlations based on errors affecting the mental representation of the choice stimuli, with the perceptual noise variance entering both the

[1] Other examples of stochastic models are Busemeyer and Townsend (1993) and Gul and Pesendorfer (2006). Both the cognitive foundations and the predictions of my model, however, are quite distinct from these models.

[2] Some recent papers have discussed related issues (Wilcox 2011, Blavatskyy 2014, Apesteguia and Ballester 2018). The insights presented here, however, rely on a model making precise predictions on the form of correlations we ought to observe in standard 'PT plus additive noise' setups based on a fully stochastic decision model.

definition of the decision parameters and the noise term. As showcased by Buschena and Zilberman (2000), this may not only affect preference parameters in quantitative terms, but can hit at the very heart of the debate on the descriptive validity of different decision models. Alós-Ferrer and Garagnani (2022) indeed showed that imposing a symmetric noise structure on data can produce inferences seemingly supporting a given model even if the underlying data have been generated based on a random choice algorithm.

The second test concerns PT's strict precept of probability-outcome separability. Violations of this precept under risk are well known, and take the form of probability weighting being affected by stake size (Hogarth and Einhorn 1990, Fehr-Duda et al. 2010, Bouchouicha and Vieider 2017). The BIM further predicts a novel violation of probability-outcome separability under ambiguity. While models of decision making typically distinguish categorically between risk and ambiguity (Ghirardato et al. 2004, Wakker 2010, Abdellaoui et al. 2011), the Bayesian inference model I present places the two on a continuum of uncertainty about choice stimuli. Knowledge about probabilities is thereby predicted to affect utility curvature, so that ambiguity will be reflected in utility curvature as well as in probability weighting. The driving force is once again complexity, since ambiguous probabilities are arguably more difficult to assess than known probabilities. Here, I test this prediction using an original dataset that is rich enough to allow for the separate estimation of all the relevant quantities under both risk and ambiguity.

Finally, I present a test pitching PT against the BIM on binary choice data, using the 'mirrors' of risky choice situations introduced by Oprea (2022). If some subjects simply adopt a heuristic of switching towards the middle of any given list, this could yield inferences that mechanically resemble likelihood-distortions (Vieider 2018). I thus compare binary choices for risk to binary choices for mirrors of the same tasks. The BIM makes no fundamental distinction between these two settings.[3] PT, on the other hand, models the first setting as being driven by attitudes towards risk, whereas the second setting reduces to a comparison of sure outcomes. I find behaviour for mirrors to closely correspond to behaviour for risky lotteries. The BIM produces comparable parameter estimates in the two settings. This illustrates the ability of the BIM to capture and quantify patterns of choice beyond the realm of risk and ambiguity.

The approach taken here builds on an influential theoretical paradigm in neuroscience, according to which the brain uses probabilistic mechanisms to encode perceptual information about the world, and decodes this information by Bayesian updating with a mental prior (Knill and Pouget 2004, Doya et al. 2007, Vilares and Kording 2011). This will be optimal whenever decision settings are complex, and thus taxing for the constrained mental resources allocated to a problem (Bhui et al.

[3] The BIM is not alone in being able to account for similar behaviour in risky choices and choices without risk maintaining the same level of complexity, see e.g. Steiner and Stewart (2016) and Netzer et al. (2021).

2021). Noise in the perception of choice quantities will then arise in quick approximate tradeoffs between multi-dimensional choice options. This serves to add computational realism to economic choices, which is often lacking in traditional models (Bossaerts et al. 2019).

The noisy neural coding of choice stimuli will naturally result in discrimination difficulties for values that are observed infrequently, an intuition that is shared with a variety of other models (Robson 2001, Netzer 2009, Woodford 2012, Steiner and Stewart 2016, Netzer et al. 2021). Such discrimination difficulties for relatively infrequent stimuli will result in regression to the mean, which can explain a number of behavioural regularities (Zhang and Maloney 2012). The degree of this phenomenon will, in turn, be governed by the relative confidence subjects have in the evidence, a phenomenon that has been documented empirically under the label of *cognitive undertainty* (Enke and Graeber 2023). Note that such a system may indeed be optimal, given that the weights determining the degree of regression to the mean are endogenously determined by the relative magnitudes of perceptual noise and environmental variation. Past experiences as enshrined in the prior then serve to rein in the worst biases that could arise from the noise, which results in efficient decision-making when cognitive resources are limited and many complex decisions have to be taken in short sequence. More in general, coding noise ought to be seen as subject to costly attentional modulation, so that it can be optimally adjusted whenever the situation warrants this.

The model I present is closely related to some recent contributions in economics. Notably, Khaw et al. (2021) used a formally similar noisy numerical perception setup to explain small stake risk aversion. My model differs from theirs in several respects. While they assume probabilities to be objectively perceived, I allow for the noisy perception of both probabilities *and* outcomes. Outcomes are furthermore assessed in a comparative cost-benefit setting, which formalizes the underlying intuition that noise arises from approximate tradeoffs. The noisy perception of probabilities makes the model applicable to uncertainty as well as risk, and even to deterministic mirrors of risky wagers (Oprea 2022). It furthermore allows for accommodating risk seeking for gains and risk aversion for losses, both of which are beyond the reach of the model of Khaw et al. (2021).

The BIM is also closely related to other recent models of noisy cognition. Gabaix and Laibson (2017) show that hyperbolic discounting can emerge from the noisy perception of future utilities by otherwise perfectly patient Bayesian decision makers. In a twin paper (Vieider 2021), I discuss a model applying an identical numerical perception setup to time delays, and show how the model accounts for several open puzzles. Natenzon (2019) uses a related Bayesian noisy perception setup to explain the attraction and compromise effects. The results I present are further related to a large literature showcasing the instability of revealed preferences. Revealed preferences ought to be unstable to the extent that individual choices are based on noisy signals, and to the extent that priors change over time. Such instability has been documented extensively for inter-temporal

correlations of risk attitudes—see Chuang and Schechter (2015) for a review. The same sort of instability has also been documented specifically for PT parameters (Zeisberger et al. 2012).

The insights I present have potentially far-reaching policy implications. For one, the model suggests that revealed choice behaviour is not indicative of actual risk preferences, making it optimal for a welfare-maximizing policy maker to ignore such patterns. It should furthermore become possible to steer behaviour by simplifying the decision environment, providing training on how to efficiently handle complex decision situations, or by providing advice to overcome such complexities. Indeed, if what we think of as preferences is driven to a large extent by the noisy processing of complex decision situations and by changeable priors, then such decisions should be eminently malleable and subject to policy intervention.

2. The Bayesian Inference Model

2.1. Noisy coding of uncertainty as log-likelihood ratios

I start by showcasing the efficiency of coding probabilistic information in terms of likelihood ratios. Assume a decision maker (DM) needs to take a decision based on her assessment of the likelihood of an uncertain event e.[4] Assume that the DM does not know the true likelihood of e occurring, but observes a noisy signal, s, about this likelihood. Assume further that the DM knows the probability of the signal conditional on the event, $P[s|e]$, as well as the probability of the signal conditional on the complementary event $\tilde{e}$, $P[s|\tilde{e}]$. The DM can then use the following equation to infer the likelihood of the event, conditional on the signal:

$$\frac{P[e|s]}{P[\tilde{e}|s]} = \frac{P[s|e]}{P[s|\tilde{e}]} \times \frac{\widehat{P}[e]}{\widehat{P}[\tilde{e}]}, \tag{1}$$

where the ratio $\frac{\widehat{P}[e]}{\widehat{P}[\tilde{e}]}$ indicates the prior likelihood ratio, which incorporates any knowledge the DM may have previously held about the likelihood of event e.[5]

This setup can be used to arrive at an optimal choice rule based on the relative costs and benefits of different actions. Take a wager offering x conditional on event e, but $y < x$ if $\tilde{e}$ obtains instead, and assume it is compared to a sure outcome, c. Integrating the costs and benefits from taking the wager into equation 1, the wager will be preferred to the sure option whenever:

$$\frac{P[e|s]}{P[\tilde{e}|s]} \times \frac{(x-c)}{(c-y)} > 1, \tag{2}$$

[4] I will generally work with discrete events (e.g., a ball extracted from an urn is blue), and associated probabilities $P[e]$. For continuous outcome variables (the stock market index increases by over 2% in a given year; cumulative rainfall during the agricultural planting season falls between 500mm and 700mm), this constitutes a slight abuse of notation for the more accurate $P[a < e < b] = \int_a^b f[e]de$, where f indicates the probability density function.

[5] This follows from an application of Bayes rule, whereby $P[e|s] = \frac{P[s|e] * \widehat{P}[e]}{P[s]}$. The use of likelihood ratios means that $P[s]$ conveniently cancels out of the expression.

where $x-c$ indicates the benefit from choosing the wager over the sure amount, and $c-y$ the cost (see Loomes 2010, for a model based on a similar intuition). This choice rule is optimal, being based on Bayesian updating and expected value maximization. It is further used without loss of generality. If the DM has a concave utility function defined over lifetime wealth, the choice model can be augmented by such a preference without altering the conclusions that follow.

From a computational point of view, it is convenient to take the natural logarithm of (2), thus transforming the choice rule from a product to a sum. Gold and Shadlen (2001) explicitly discuss the neural realism of such a choice rule. The optimality of the choice rule is unaffected by such an operation, since any monotonic transformation will leave this choice rule unaltered. This holds both ex ante *and* ex post, and the results presented below do not change in any substantive way if I use the choice rule in (2) instead (see appendix S1). The choice rule then becomes:

$$ln\left(\frac{P[e|s]}{P[\tilde{e}|s]}\right) > ln\left(\frac{c-y}{x-c}\right). \tag{3}$$

The wager should be chosen over the sure option whenever the log-likelihood ratio of event e and its complement $\tilde{e}$ exceeds the log cost-benefit ratio.

The model can easily be extended to the comparison of multi-outcome wagers. Take two wagers offering outcomes $\boldsymbol{x}=\{x_1,...,x_n\}$ and $\boldsymbol{y}=\{y_1,...,y_n\}$ under events $e_1,...,e_n$, where wager $\boldsymbol{x}$ is riskier. The event space is constructed in a way as to encompass all comparative events. The optimal choice rule equivalent to (2) then takes the following form:

$$\sum_{i=1}^{n-1}\frac{P[e_i]}{P[e_n]}\frac{(x_i-y_i)}{(y_i-x_n)} > 1, \tag{4}$$

which sums the relative costs and benefits of the two wagers. The equation takes the form of an optimal choice rule used in signal detection theory (Green et al. 1966). It incorporates the same two principles as (2)—optimal belief updating, and assuming additivity of probabilities, expected value maximization. The reference outcome, here taken to be x_n, is arbitrary, since the choice rule enshrines within it all pairwise comparisons (similar to what happens in multinomial logits). To see this, let $V_{in} \triangleq \frac{P[e_i]}{P[e_n]}\frac{(x_i-y_i)}{(y_i-x_n)}$. It is then straightforward to derive any binary comparison as $V_{ij}={}^{V_{in}}\!/_{V_{jn}}$.

The model I will present below is based on the premise that mental processing of uncertain wagers functions much like in the optimal choice rule above. According to an influential theoretical paradigm in neuroscience, the human brain acts as a Bayesian inference machine, continuously combining noisy signals about the environment with prior beliefs to assess the relative merit the choice options (Knill and Pouget 2004, Doya et al. 2007, Vilares and Kording 2011). Even numerically represented quantities, such as a monetary outcome x or an objectively given probability p,

will be mentally represented by a noisy signal before entering the choice process. Noise will arise especially when assessments of the relevant quantities are made quickly and intuitively. Such noise may indeed be an optimal reaction to many complex tradeoffs being presented in short sequence. In general, however, noise ought to be seen as subject to attentional modulation, so that it will be a function of the importance and the complexity of the decision at hand.

2.2. Bayesian mental processing of noisy signals

I now present a step-by-step derivation of subjective choice parameters from the mental encoding-decoding of choice stimuli. For simplicity's sake, I will present the derivation for the choice between a binary wager and a sure outcome, as in (3). The derivation for the more general choice rule in (4) proceeds very similarly, and is shown in Appendix A.

The mental encoding stage and the likelihood

Take a vector of mental signals, $\boldsymbol{r}$, encoding the characteristics of a given choice problem. The signals in this vector can be thought of as neural firing rates, encoding the desirability of the choice stimuli in the brain. The mental signals in $\boldsymbol{r}$ thus play a role that is analogous to that of the signal s used above to illustrate the optimal choice rule. Just like s above, the mental signals will generally be noisy, and the noise itself may carry useful information about the accuracy of the signals. Other than above, however, not only uncertain events are represented by mental signals, but so are nominally certain monetary outcomes, the exact nature of which also needs to be inferred by the mind from noisy signals. I will thus assume that there are two such mental signals in $\boldsymbol{r}$, r_e and r_o, encoding the desirability of the likelihood ratio and the cost-benefit ratio, respectively.[6] Costs $c-y$ and benefits $x-c$ from taking the wager are best conceived of as approximate comparisons, rather than the precise mathematical quantities entailed by the differences depicted, thus providing a further rationale for their noisy mental representations.[7]

Stimuli will have compressed mental representations for efficiency reasons, and logarithmic functions are typically used to model such mental representations (Dayan and Abbott 2001, Petzschner et al. 2015). It has been shown that the activation of neurons responsible for the approximate mental representations of numbers is best fit by a normal distribution with as its argument the

[6] The setup with two mental signals corresponds to a minimal setup in which the two choices can be compared. In particular, modelling costs and benefits through two separate mental signals will only result in a rescaling of decision noise, and is thus inconsequential from any practical point of view. I abstract from explicitly modelling the encoding of other aspects of the decision situation, such as the colours associated with winning and losing options, since they are not of central importance to the decision to be taken and do not enter the optimal choice rule derived above.

[7] In particular, transformation of single numerical quantities such as modelled by Khaw et al. (2021) could be included in the equation to provide a further rationale for the noisy comparison modelled here, so that my approach is complementary to theirs. Tversky (1969) discusses the superiority of such a comparative setup, inasmuch as it contains the transformation of single quantities within as a special case and facilitates otherwise taxing value judgments. I refrain from formally introducing such additional transformations for the sake of analytical tractability.

logarithm of the number being represented (Dehaene and Changeux 1993, Dehaene 2003, Nieder and Miller 2003, Piazza et al. 2004, Harvey et al. 2013). The underlying rationale for logarithmic coding can be traced back to the functional architecture of the neural system, where single neurons tend to have receptive fields that are tuned towards the detection of stimuli falling into a pre-determined range. Howard and Shankar (2018) show that logarithmic spacing of receptive fields is optimal inasmuch as it allows an organism to adapt to the statistics of the environment with maximum flexibility. This conclusion holds independently of the distribution of actual stimuli in the environment, which is a priori unknown to the organism.

I thus assume that r_e and r_o are two independent signals representing the log-likelihood ratio and the cost-benefit ratio. Note that r_e and r_o constitute a shorthand for the more accurate $r_e|e$ and $r_o|x,y,c$, since the signals are conditional on the specific stimulus being presented. The term $ln\left(\frac{P[e]}{P[\tilde{e}]}\right)$ thus indicates the mental probability representation in the expression below, which is assumed to correspond to the true underlying probabilities in the case of fully described risks. The same holds for Ellsberg urns, where the true probability can be inferred from an exchangeability argument. The two signals are independently drawn from the following likelihood functions:

$$r_e \sim \mathcal{N}\left(ln\left(\frac{P[e]}{P[\tilde{e}]}\right), \nu^2\right) \quad , \quad r_o \sim \mathcal{N}\left(ln\left(\frac{c-y}{x-c}\right), \nu^2\right), \tag{5}$$

where ν is the common coding noise of the log odds and the log cost-benefit ratio, capturing the uncertainty around the true mean of the stimulus with which a given stimulus is perceived.[8] The coding noise is modelled as being common to the two dimensions since different coding noises would not be separately identifiable from direct tradeoffs—an important feature of a model that is supposed to be neurally implemented. This same assumption is indeed maintained throughout the Bayesian noisy coding literature (Natenzon 2019, Khaw et al. 2021). The representation as the logarithm of the stimuli implies that the difference necessary for two stimuli to be reliably discriminated will be proportional to the magnitude of the stimuli themselves (Dayan and Abbott 2001, ch. 3). This amounts to a so-called *just noticeable difference* in the stimulus m, Δm, so that $\frac{\Delta m}{m}$ is a constant. Neural coding can thus be seen as the driving factor behind a behavioural phenomenon that has long been known in psychophysics as the Weber-Fechner law (Fechner 1860).

The mental decoding stage and the posterior expectations

The information provided by the noisy mental signals r_e and r_o needs to be decoded to be transformed into actionable quantities that can inform the decision process. This is due to the uncertainty in the mental representation of the stimuli, which makes it desirable to combine the encoded

[8] In the case of degenerate choices involving e.g. $x = c$ or $c = y$, the equations in (5) are not defined. This can easily be fixed by adding an arbitrarily small number, ϵ, to both the numerator and denominator inside the logarithm. I do not do this here to avoid notational clutter.

signal with prior information about what sort of stimuli are likely to be encountered in the given decision environment. It seems natural to let the mental priors used to this effect follow a normal distribution, since the choice of a conjugate prior distribution will minimize the burden in terms of neural computations. This, once again, serves to increase the biological realism of the model. The priors take the following form:

$$ln\left(\frac{P[e]}{P[\tilde{e}]}\right) \sim \mathcal{N}\left(ln(\xi)\,,\,\sigma_e^2\right) \;\;,\;\; ln\left(\frac{c-y}{x-c}\right) \sim \mathcal{N}\left(ln(\zeta)\,,\,\sigma_o^2\right) \tag{6}$$

where $ln\,(\xi)$ and $ln(\zeta)$ are the means and σ_e and σ_o are the standard deviations of the two priors. The logit-normal distributions imply that the prior mean will be 0 for likelihoods and costs and benefits that are equal on average, and will tend towards plus and minus infinity as the differences between these quantities become extreme. Such symmetric distributions around the mean provide a realistic representation of typical choice quantities a DM would encounter across a range of environments.

Combining the likelihoods in (5) with the priors in (6) by Bayesian updating, we obtain the following posterior distributions, *conditional on the mental signals drawn*:

$$\begin{aligned} ln\left(\frac{P[e]}{P[\tilde{e}]}\right) \,|\, r_e &\sim \mathcal{N}\left(\frac{\sigma_e^2}{\sigma_e^2+\nu^2}\times r_e + \frac{\nu^2}{\sigma_e^2+\nu^2}\times ln\,(\xi)\;,\;\frac{\nu^2\sigma_e^2}{\nu^2+\sigma_e^2}\right) \\ ln\left(\frac{c-y}{x-c}\right) \,|\, r_o &\sim \mathcal{N}\left(\frac{\sigma_o^2}{\sigma_o^2+\nu^2}\times r_o + \frac{\nu^2}{\sigma_o^2+\nu^2}\times ln\,(\zeta)\;,\;\frac{\nu^2\sigma_o^2}{\nu^2+\sigma_o^2}\right), \end{aligned} \tag{7}$$

where $\frac{\sigma_e^2}{\sigma_e^2+\nu^2}$ and $\frac{\sigma_o^2}{\sigma_o^2+\nu^2}$ constitute the Bayesian shrinkage weights assigned to the signals of the log-likelihood ratio and the log cost-benefit ratio of the stimulus, relative to the weights assigned to the means of the priors. Given the importance of these weights for what follows, I will henceforth use the following shorthand for the weights:

$$\gamma \triangleq \frac{\sigma_e^2}{\sigma_e^2+\nu^2} \;\;,\;\; \alpha \triangleq \frac{\sigma_o^2}{\sigma_o^2+\nu^2}, \tag{8}$$

so that $\frac{\nu^2}{\sigma_e^2+\nu^2} = 1-\gamma$ and $\frac{\nu^2}{\sigma_o^2+\nu^2} = 1-\alpha$. The relative uncertainty associated with the mental signal and the prior, as captured by ν and $\{\sigma_e, \sigma_o\}$, will thus determine how much weight will be attributed to the signal versus the prior in the posterior mean. As the coding noise parameter ν converges to 0, the signal will accurately reflect the stimuli, and the weights put on the signals will converge to 1. For coding noise strictly larger than 0 and weights strictly smaller than 1, however, the expectations of the log-likelihood ratio and the log-cost benefit ratio conditional on the signals will reflect a convex combination of signals and priors.

The actionable choice rule

The posterior expectations of the choice quantities just derived can now inform the decision process.

To this end, we must amend the optimal choice rule described in equation 3 by replacing the objective quantities with their mental representations:

$$E\left[ln\left(\frac{P[e]}{P[\tilde{e}]}\right)|\, r_e\right] > E\left[ln\left(\frac{c-y}{x-c}\right)|\, r_o\right], \tag{9}$$

indicating that the wager on event e will be accepted whenever the posterior expectation of the log-likelihood ratio exceeds the posterior expectation of the log cost-benefit ratio. Note, however, that the choice rule will nevertheless not be deterministic, since it depends on the stochastic signals r_e and r_o, which are drawn from the distributions in (5). Substituting the posterior expectation from (7) into the choice rule in (9) and solving for the mental signals, we get:

$$\gamma \times r_e - \alpha \times r_o > ln(\delta)^{-1}, \tag{10}$$

where $\delta \triangleq \xi^{1-\gamma} \times \zeta^{\alpha-1}$, and where $ln(\delta)^{-1}$ provides the threshold which the weighted difference of mental signals on the left-hand side needs to exceed in order for the wager to be accepted. The threshold parameter δ has a natural interpretation, since it is made up by the prior mean of the likelihood ratio, multiplied by the prior mean of the benefit-cost ratio (i.e., the inverse of the cost-benefit ratio)—the more favourable the prior expectation of the likelihood ratio and of the benefit to cost ratio, the more likely the DM will be to accept the wager.

The probabilistic choice rule

To derive an expression for the probability with which the wager will be chosen over the sure outcome that can be observed by an econometrician, we need to obtain an expression free of the unobservable mental signals. To this end, we obtain the z-score of the weighted difference in mental signals in (10) by jointly distributing the two signals exploiting the known distributions in (5). Obtaining the z-score, and comparing it to the z-score of the threshold equation yields the following probabilistic choice rule (see proof below):

$$Pr[(x,e;y) \succ c] = \Phi\left(\frac{\gamma \times ln\left(\frac{P[e]}{P[\tilde{e}]}\right) - \alpha \times ln\left(\frac{c-y}{x-c}\right) - ln(\delta)^{-1}}{\nu \times \sqrt{\gamma^2+\alpha^2}}\right), \tag{11}$$

where Φ is the standard normal cumulative distribution function. Notice how the same parameters governing the discriminability between the choice options are also driving decision noise. All model parameters are indeed tightly intertwined, with coding noise ν constituting the common cause binding everything together. The noise term in the denominator can further be rewritten as $\sqrt{\frac{\nu^2\sigma_e^4}{(\nu^2+\sigma_e^2)^2} + \frac{\nu^2\sigma_o^4}{(\nu^2+\sigma_o^2)^2}}$, showing that it is made up of the variances of the *response distributions*, i.e. of the distributions of the inferred choice stimuli conditional on the real choice stimuli (see Ma et al. 2023, ch. 4). Stochasticity in choices is thus driven by the average uncertainty about the posterior inferences on the choice stimuli.

Proof of Probit equation. Given that r_e and r_o follow a normal distribution, their weighted difference in (10) will itself follow a normal distribution:

$$\gamma \times r_e - \alpha \times r_o \sim \mathcal{N}\left(\gamma \times ln\left(\frac{P[e]}{P[\tilde{e}]}\right) - \alpha \times ln\left(\frac{c-y}{x-c}\right), \omega^2\right),$$

where $\omega \triangleq \nu \times \sqrt{\gamma^2 + \alpha^2}$ is the standard deviation of the weighted difference of the two uncorrelated noisy mental signals. This yields the following z-score:

$$z = \frac{\gamma \times r_e - \alpha \times r_o - \left[\gamma \times ln\left(\frac{P[e]}{P[\tilde{e}]}\right) - \alpha \times ln\left(\frac{c-y}{x-c}\right)\right]}{\omega}.$$

Transforming the threshold equation (10) to be expressed on the same scale, we obtain

$$z_t = \frac{\gamma \times r_e - \alpha \times r_o + ln(\delta)^{-1}}{\omega}.$$

Subtracting $z - z_t$ eliminates the unobservable signals and yields the expression inside parentheses in (11). Since a z-score follows a standard-normal distribution by definition, a difference between two z-scores will also follow a standard normal distribution. □

2.3. The BIM generates PT functionals for binary wagers

I now show that the BIM results in stochastic micro-foundations for choice patterns as they have been documented in the PT literature. We can write the point of indifference in the numerator in (11) as follows, where indifference is taken to mean a 50% chance of choosing either choice option:

$$ln\left(\frac{c-y}{x-c}\right) = \alpha^{-1}\left[ln(\delta) + \gamma \times ln\left(\frac{P[e]}{P[\tilde{e}]}\right)\right]. \tag{12}$$

This represents the point of indifference between $(x, e; y)$ and c, where the likelihood dimension on the right—subjectively transformed by the mental representation parameters α, γ, and δ—is traded off against the outcome dimension on the left. Let us define $\pi(P[e]) \triangleq \frac{c-y}{x-y}$, which is the solution of the equation $c = \pi(P[e])x + (1-\pi(P[e]))y$ for the decision weight $\pi(P[e])$ under a dual-EU representation of the choice problem (Yaari 1987). Assuming $P[\tilde{e}] = 1 - P[e]$ following PT, we obtain $ln\left(\frac{c-y}{x-c}\right) = ln\left(\frac{\pi(P[e])}{1-\pi(P[e])}\right)$. Substituting this expression into (12) and solving for $\pi(P[e])$ yields:

$$\pi(P[e]) = \frac{\delta^{1/\alpha}(P[e])^{\gamma/\alpha}}{\delta^{1/\alpha}(P[e])^{\gamma/\alpha} + (1-P[e])^{\gamma/\alpha}}. \tag{13}$$

This closely resembles a probability-distortion function commonly used in the decision-making literature (Goldstein and Einhorn 1987, Gonzalez and Wu 1999, Bruhin et al. 2010). The variance ratio $\alpha \leq 1$ allows for outcome distortions in addition to probability distortions, and shows how probability weighting emerges from the tension between the two. The parameter γ mostly governs the slope of the function, capturing likelihood-sensitivity. The fact that γ decreases in the noisiness

of the coding process receives support from findings showing that probabilistic sensitivity increases with cognitive ability (Choi et al. 2022). The parameter δ determines mostly the elevation of the function, thus governing global optimism/pessimism.

The mapping just presented shows how PT-like parameters naturally emerge from noisy mental representations of choice stimuli and their mental decoding by a prior distribution, indicating the likelihood of different stimuli in a given environment. A difference from PT is that outcome-distortions are defined over the costs and benefits associated with different events, rather than over single outcomes. Such a comparative setting has a number of advantages over outcomes being transformed individually. For one it is more general, and could accommodate transformations of single outcomes within it (Tversky 1969). It will also facilitate the evaluation of choice tasks, since outcomes are compared directly instead of being evaluated separately. The comparative setting thus seems well-suited for the approximate valuation entailed by the BIM, since it allows to simplify the problem when some quantities are approximately equal.

We can also conceive of x, y, and c as losses, with $x > c > y$. The setup just derived can then directly be applied to this situation. Assume that under event e the DM stands to lose x, or else lose $y < x$ under $\tilde{e}$, and that this scenario is compared to a sure intermediate loss of c. The gains and losses are now flipped relatively to the setup discussed above, so that $x - c$ constitutes the cost from taking the wager, and $c - y$ the benefit, with everything else remaining the same. Such a setup will then naturally result in decreasing sensitivity towards costs and benefits in both the gain and loss domain, much like PT incorporates decreasing sensitivity towards absolute outcomes relative to a reference point. Furthermore, the mechanism used to account for loss aversion in Khaw et al. (2021) carries over to this setting, so that the model can also account for loss aversion.[9]

We can further derive substantive predictions about behaviour from the intuitions emerging from the model. Under the BIM, the parameter δ has a natural interpretation as the mean of the prior. Let $\xi = \frac{\psi}{1-\psi}$ and $\alpha = 1$, so that $\delta = \left(\frac{\psi}{1-\psi}\right)^{1-\gamma}$. The parameter ψ can then be shown to coincide with the fixed point of the probability-distortion function where the function crosses the 45° line. Substituting ψ into the decision weight for the probability $P[e]$, we obtain:

$$ln\left(\frac{\pi(\psi)}{1-\pi(\psi)}\right) = \gamma \times ln\left(\frac{\psi}{1-\psi}\right) + (1-\gamma) \times ln\left(\frac{\psi}{1-\psi}\right) = ln\left(\frac{\psi}{1-\psi}\right), \tag{14}$$

from which follows that $\pi(\psi) = \psi$, and by extension, $\pi(P[e]) = P[e]$. The expectation of the mental prior thus coincides with the fixed point of the probability-distortion function, at which probabilities are perceived without subjective distortions.

[9] Khaw et al. (2021) describe loss aversion by different priors for gains and losses. If the prior mean for losses is larger in absolute value than the one for gains, then loss aversion takes the form of a multiplicative constant just like in PT under some additional constraints on the variances.

Decisions under risk and ambiguity. The BIM can directly be applied to decision-making under risk and ambiguity. The case of risk obtains by setting $P[e] = p$, with p a fully described probability. The case of ambiguity obtains when subjects are asked to bet on Ellsberg-urns with unknown colour proportions (Ellsberg 1961, Abdellaoui et al. 2011). Given that unknown odds are more difficult to encode than known odds (Petzschner et al. 2015), we should expect coding noise to increase as the information about the probabilities involved becomes more vague.[10] This yields the prediction that $\gamma_a < \gamma_r$, where the subscripts a and r stand for 'ambiguity' and 'risk' respectively—a phenomenon known as *ambiguity-insensitivity*, which is well-documented (Abdellaoui et al. 2011, Dimmock et al. 2015, Trautmann and van de Kuilen 2015, L'Haridon et al. 2018). Once again, this prediction is driven by the Bayesian aggregation of evidence and prior. With the evidence carrying less weight under ambiguity, we should expect regression to the mean of the prior to increase in strength. This prediction is further supported by the finding that time pressure, which presumably augments coding noise, increases ambiguity-insensitivity (Baillon et al. 2018b).

It is now at the time to circle back to the noise term. Other than under typical stochastic implementations of PT, where the decision model and the noise model are combined in an *ad hoc* fashion, decision noise naturally emerges from the same mental encoding-decoding process as the other model parameters. Outcome-discriminability α and likelihood-discriminability γ directly enter the definition of decision noise, $\omega \triangleq \nu \times \sqrt{\alpha^2 + \gamma^2}$, and coding noise enters the definition of the discriminability parameters $\alpha \triangleq \frac{\sigma_o^2}{\sigma_o^2+\nu^2}$ and $\gamma \triangleq \frac{\sigma_e^2}{\sigma_e^2+\nu^2}$. This implies that both the choice parameters and decision noise will be causally determined by noise in the signals as quantified by ν. It follows that parameters estimated in a PT setup, which neglects these intricate interrelations, should be correlated in a systematic way. This concerns primarily correlations between the noise parameter and likelihood-sensitivity, with knock-on effects on other parameters. (Being defined over costs and benefits, the outcome-discriminability parameter, α, deviates more substantially from utility curvature in PT, so that the predictions are less crisp). The correlations are predicted to occur because of an 'omitted variable bias' in the 'PT plus white noise' setup, with coding noise ν playing the role of the omitted variable according to the predictions of the BIM.

Coding noise and separability violations. A further consequence of the role played by ν in the definitions of α and γ is that no strict separability between the likelihood and outcome dimensions seems warranted, in opposition to what is postulated by PT. We may thus expect coding noise to increase in stakes. This should then be reflected in lower outcome-discriminability,

[10] I limit my discussion to the case of Ellsberg urns with unknown colour proportions. Findings may well deviate from the ones described here in other contexts, since it is well-known that ambiguity attitudes over natural sources of uncertainty may vary depending on a DM's knowledge of the specific context or the DM's perceived competence in a given task (Heath and Tversky 1991, Abdellaoui et al. 2011). Presumably, such effects may be created by learning (Baillon et al. 2018a), but a formal treatment of this issue is beyond the scope of this paper.

resulting in apparent patterns of increasing relative risk aversion (Holt and Laury 2002). The same change will also result in apparent changes in likelihood-discriminability, thus impacting the probability dimension. Under PT, such effects will take the form of violations of probability-outcome separability, whereby changes in stakes ought to be reflected purely in utility curvature and leave probability-distortions unaffected. Such violations are well-documented in the literature (Hogarth and Einhorn 1990, Fehr-Duda et al. 2010, Bouchouicha and Vieider 2017).

The flip-side of this issue is observed when moving from risk to ambiguity. Given the increase in coding noise we expect under ambiguity, the BIM predicts a lowered level of likelihood-discriminability. The latter, in turn, is expected to go hand-in-hand with a lowering of outcome-discriminability, which results from the increase in coding noise for equal variance of choice stimuli, so that σ_e can be expected to remain constant across the two settings. While remaining undocumented to date, such a pattern would contradict PT, according to which ambiguity attitudes ought to be reflected purely in probability weighting (Wakker 2010, Abdellaoui et al. 2011, Dimmock et al. 2015).

3. Empirical evidence

3.1. PT parameters show systematic correlations with noise

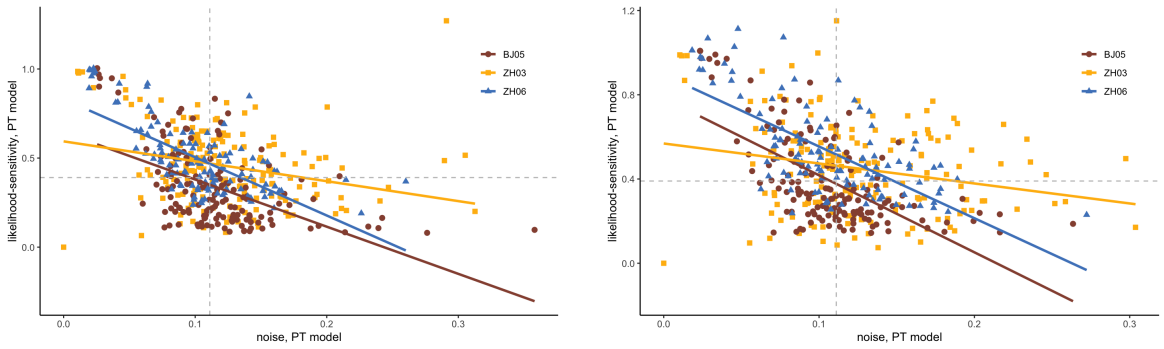
I start by examining parameter correlations in PT models. I use the rich individual-level data of Bruhin et al. (2010) collected in 3 different experiments, to estimate distributions of PT parameters (results based on the 30 country data of L'Haridon and Vieider 2019 are similar, see appendix S2.2). The basic econometric framework follows the one in the original paper (see appendix S2.1 for details). Other than in the original paper, I use Bayesian random parameter models to estimate the model parameters (Gelman et al. 2014b). This method is geared towards maximizing the *predictive* power of the model for new data, rather than towards optimizing its fit to existing data.

The BIM predicts that the parameters governing sensitivity towards outcomes and probabilities are governed by coding noise. This results in a prediction that, when estimating a PT model, the noise term ought to be correlated with the decision parameters, given that the residuals will also be determined by coding noise. In a typical PT setup plus noise, on the other hand, the additive noise term is assumed to take the form of 'white noise', and thus to be independent of the decision model itself. The same issues, however, also generalize to more sophisticated noise structures such as the contextual utility model of Wilcox (2011), since they do not take the specific mechanism predicted by the BIM into account. A violation of independence of preference parameters and noise would be problematic, since any inferences on preference parameters one draws from the model may then be sensitive to the specific assumptions about noise. I will test these predictions using Spearman rank correlations on the estimated parameters.[11] Any p-values reported are always two-sided.

[11] I use rank correlations to minimize the impact of potential outliers. Correlations based on the estimated covariance matrix or calculated using the full richness of the Bayesian posterior draws deliver analogous results.

To ensure comparability with the model derived above, I adopt power utility throughout, so that $u(x) = x^{\widehat{\alpha}}$. I will use the Goldstein and Einhorn (1987) probability-distortion function for the same reason, with parameters $\widehat{\gamma}$ and $\widehat{\delta}$, where the 'hat' serves to distinguish the PT parameters from the equivalent BIM-generated parameters. These same functional forms were also used in the original paper. I will refer to $\widehat{\gamma}$ as *likelihood-sensitivity*, and to $\widehat{\alpha}$ as *outcome-sensitivity*, to distinguish them from the equivalent BIM-derived parameters, to which I refer as *discriminability* parameters.

Figure 1 shows correlations between decision noise and likelihood-sensitivity $\widehat{\gamma}$ for gains and for losses. Note that the econometric model allows for heteroscedasticity in errors across outcome domains, and the errors for gain and losses are thus not the same. All results presented are robust to using contextual errors as proposed by Wilcox (2011) instead. Panel 1(a) shows a scatter plot of the PT noise parameter, $\widehat{\omega}^{+}$, against the likelihood-sensitivity parameter, $\widehat{\gamma}^{+}$, for gains. The two parameters show a strong negative correlation ($\rho = -0.422, p < 0.001$), which is also present in each of the 3 individual experiments. The results for losses, shown in panel 1(b), are very similar ($\rho = -0.389, p < 0.001$). For both gains and losses, a small group stands out that has virtually no noise and likelihood-sensitivity arbitrarily close to 1. These are the expected value maximizers detected in the mixture model of Bruhin et al. (2010), who most likely based their responses on precise calculations rather than on quick and approximate judgments. The correlations presented are qualitatively robust to removing this group of subjects.



(a) Scatter plot of $\widehat{\omega}^{+}$ and $\widehat{\gamma}^{+}$ (b) Scatter plot of $\widehat{\omega}^{-}$ and $\widehat{\gamma}^{-}$

Figure 1 Scatter plot of PT parameters, likelihood-sensitivity

The parameters have been obtained from the estimation of a PT model plus additive noise. The different colours and shapes represent the 3 experiments in Bruhin et al. (2010): ZH03 stands for Zurich 03; ZH6 for Zurich 06; and BJ05 for Beijing 05. Solid lines indicate regression lines. The dashed lines indicate the median parameter values. Some outliers may be cut for better visual display.

The correlations just shown further have knock-on effects on correlations between the deterministic model parameters. Larger values of $\widehat{\gamma}$ coincide with values of $\widehat{\delta}$ tending towards 1 (see appendix S2.2). These effects are not foreseen by PT, but line up exactly with the predictions

emerging from the BIM. There are furthermore correlations between likelihood-sensitivity $\widehat{\gamma}$ and outcome-sensitivity $\widehat{\alpha}$, even though the predictions of the BIM are less crisp in this respect due to the different definitions of outcome transformations in the two models (see Herold and Netzer 2023, for predictions on such correlations). Online appendix S2.3 further reports predictive tests of the BIM against PT, and shows that the BIM outperforms 'PT plus white noise' (as well as PT plus contextual noise) in all 6 test cases of the Bruhin et al. (2010) data, as well as in the 60 test cases provided by the L'Haridon and Vieider (2019) data.

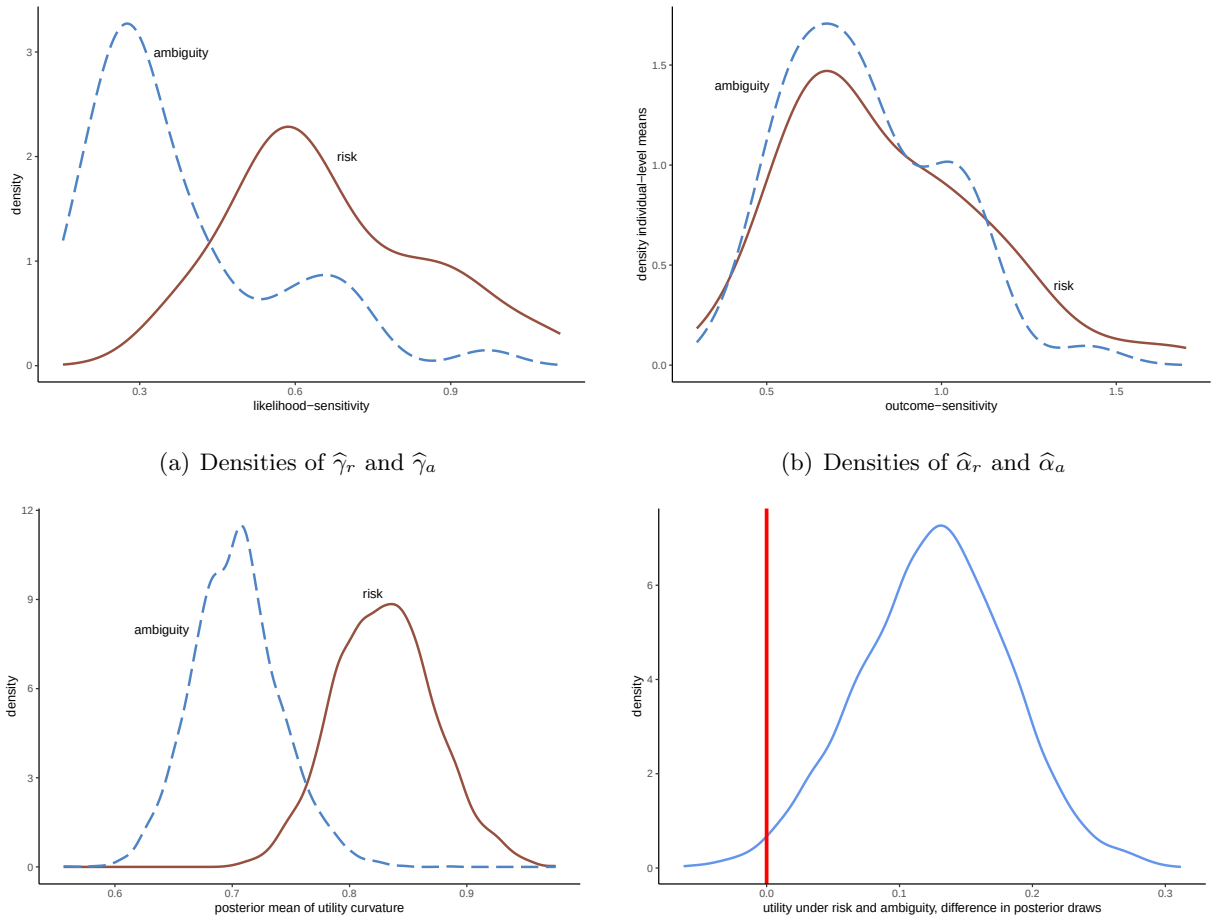
The patterns I just illustrated are clearly suggestive of the patterns predicted by the BIM. In light of the extensive empirical literature estimating PT functionals, it seems surprising that such correlations have not been examined previously. The reason for this may be that no existing theoretical models—including existing stochastic choice models—have predicted such correlations. Nevertheless, this does not exclude that some alternative error model *could* be devised to account for such patterns, though such a model would likely need to capture some of the same elements enshrined in the BIM.

3.2. Probability-outcome separability under ambiguity

The BIM predicts ambiguity-insensitivity based on increased difficulty in assessing unknown probabilities. Such patterns are well-documented in the literature (Abdellaoui et al. 2011, Dimmock et al. 2015, Trautmann and van de Kuilen 2015, L'Haridon et al. 2018). Enke and Graeber (2023) have documented empirically that this phenomenon goes hand-in-hand with a lowering of the confidence people declare to have in their choices, which is consistent with the account presented here. In addition, however, the BIM predicts a novel violation of PT's probability-outcome separability, whereby utility curvature is also affected by ambiguity.[12] This constitutes a violation of PT, since utility curvature for risk and ambiguity is assumed to be the same under PT due to the strict separation of the likelihood and outcome dimensions inherent in the model (Wakker 2010, Abdellaoui et al. 2011, Dimmock et al. 2015). In this section, I present a test of that prediction in a PT setup.

I use an original dataset containing a rich choice setup to test this hypothesis. The data contain observations for 47 subjects indicating certainty equivalents for both risky and ambiguous lotteries in a within-subject design. The data structure and experimental procedures closely follow those in L'Haridon et al. (2018). The stimuli are, however, richer in that the stimuli for risk are replicated exactly for ambiguity, including variation over outcomes and non-zero lower outcomes for both risk and ambiguity, thus allowing for the identification of utility curvature under ambiguity as well as under risk (Fehr-Duda and Epper 2012). The econometric approach closely follows the setup used above for the data of Bruhin et al. (2010). Further details can be found in appendix S3.

[12] This violation arises due to an increase in coding noise under ambiguity, triggered by the increased complexity inherent to the description of Ellsberg urns. At the same time, the outcomes remain fully observable and are identical to those presented under risk, so that there is no reason for the variance of the outcome prior to change.



(a) Densities of $\widehat{\gamma}_r$ and $\widehat{\gamma}_a$

(b) Densities of $\widehat{\alpha}_r$ and $\widehat{\alpha}_a$

(c) Posterior densities of the means

(d) Density difference of means

Figure 2 Likelihood- and outcome-sensitivity under risk and ambiguity

Densities of individual-level parameters (top) and of the mean parameters (bottom) of likelihood- and outcome-sensitivity. Panel 2(d) plots the difference in posterior draws between the two means, which is the quantity used in Bayesian tests of mean differences (Kruschke 2014).

Figure 2 shows density plots for $\widehat{\gamma}$ and $\widehat{\alpha}$, for risk and ambiguity. The difference in individual-level estimates between risk and ambiguity for $\widehat{\gamma}$, shown in panel 2(a), is sizeable. This sort of lowering of likelihood-sensitivity under ambiguity relative to risk, termed ambiguity-insensitivity, is indeed a standard finding in the PT literature. The difference in individual-level parameters between outcome sensitivity for risk and ambiguity is displayed in panel 2(b). While it is less pronounced, the distribution under ambiguity is shifted to the left. This is confirmed by a Mann-Whitney test on the paired individual-level parameters, indicating that outcome-sensitivity is indeed reduced under ambiguity ($p = 0.002$, two-sided).

One can further test for the existence of differences using the posterior draws of the mean parameters for risk and ambiguity. Panel 2(c) shows the density of the posterior draws of the means of the two utility parameters for risk and ambiguity. The utility parameter for ambiguity is

clearly smaller than the one for risk. To assess the statistical significance of this difference, however, we need to look at the *difference* in posterior draws, given that the draws are not independent (Gelman et al. 2014b, Kruschke 2014). That difference is shown in panel 2(d). The overlap with 0 is minimal, and 0.987 % of the posterior probability mass falls above 0, indicating a significant difference in the means. This results in a violation of the probability-outcome separability precept of PT. It does, however, conform precisely to the prediction emerging from the BIM.

3.3. Lottery choice versus deterministic mirrors

The tests shown so far aimed at testing predictions about what we should observe if we estimated the data in a PT setup, if the BIM is the true data-generating model. Here, I present a test of PT against the number-perception account of the BIM. To this end, I deploy a variation of Oprea's (2022) deterministic mirrors. One potential shortcoming of the approach taken by Oprea (2022) is that the choice list format may be driving some of the results. That is, if some subjects pick switching points that mechanically fall towards the middle of the list, this could produce the same sort of patterns for risk and deterministic mirrors. I thus adapt the mirror setup to a rich binary choice setting to test the robustness of the effect. This allows me to structurally estimate the BIM for both risk and mirrors, which serves to showcase the BIM's ability to capture these patterns.

The experiment was conducted in an introductory class of Behavioural Economics at Ghent University using a between-subject design. Students had heard about expected value and expected utility theory, but had not yet been introduced to behavioural models of decision making. Students were told to bring a laptop or tablet to class. The lecturer seated them apart so that they could not look into each others' screens. All subjects who showed up for class participated in the experiment. While calculations of expected values on the computers were possible in principle, such calculations would work *against* the effects shown below. 10 subjects were randomly selected to play one of their decisions for real money. After a short introduction describing the length of the experiment and the incentive structure, the lecturer shared a link for the online experiment with the students. 133 students submitted complete responses in the allocated time.

Figure 3 shows a screenshot of the experiment. Subjects had to indicate their choice between two options. Each option had a number of boxes associated with monetary payoffs. Option A had some boxes associated with a higher payoff, and some with a lower payoff. The number of high outcome boxes could take the value 12, 32, 42, 48, 52, 58, 68, 88. The winning outcome ranged from €18 to €32, and the tasks included non-zero lower outcomes as well (see table 1 for a list). Option B always contained 100 boxes with the same outcome, which was made to vary in steps of €1 between the high and the low outcome in option A. The binary choice options were, however, completely randomized, as was the position of option A and option B on the screen. These stimuli were chosen

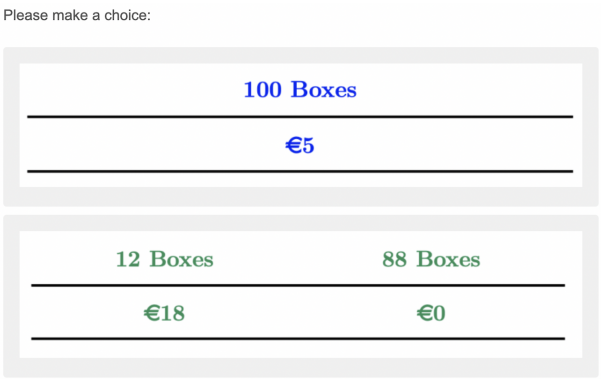


Figure 3 Screenshot of choice situation

Table 1 Choice tasks and choice proportions by treatment

	characteristics			risk		mirrors		ranksum test
task	high	low	prob	risky	SD	complex	SD	p-value
1	18	0	0.12	0.14	0.35	0.14	0.34	0.672
2	18	0	0.32	0.22	0.42	0.24	0.43	0.614
3	18	0	0.68	0.52	0.50	0.55	0.50	0.142
4	18	0	0.88	0.70	0.46	0.72	0.45	0.301
5	27	5	0.48	0.39	0.49	0.40	0.49	0.704
6	27	5	0.52	0.49	0.50	0.53	0.50	0.287
7	32	5	0.42	0.39	0.49	0.38	0.49	0.718
8	32	0	0.58	0.42	0.49	0.46	0.50	0.334
9	43	18	0.48	0.42	0.49	0.45	0.50	0.410

Summary table of task characteristics and choice proportions. The column 'risky' indicates the choice proportion of the risky option A in the random box treatment. The column 'complex' indicates the choice proportion of the complex option A in the average box treatment. The column 'ranksum test' indicates p-values of Wilcoxon rank sum tests executed on choice proportions of Option A. The tests are based on 72 subjects assigned to the average box treatment, versus 61 subjects assigned to the random box treatment.

to reflect typical choice stimuli used to investigate decision making under risk. Subjects faced 200 such binary choices in total, of which 20 were repeated choices.

The instructions provided the following information. First, all subjects saw an explanation of the general structure of the choice task. They were told that they had to decide between two options, and how different boxes could have different payoffs attached to them. Subsequently, subjects were randomly allocated to one of two treatments—a *random box* treatment, and an *average box* treatment. In the random box treatment, they learned that one random box from the chosen option would be opened to determine their payoff, following standard procedures in the risky choice literature. In the average box treatment, on the other hand, they learned that they would be paid the amount in the average box. That is, the payoffs would be summed over the 100 boxes, and divided by 100. Examples served to clearly illustrate the choice situation and resulting payoffs.

Subjects furthermore had to go through a number of comprehension questions that were meant to drive home the payoff mechanism (full instructions in appendix S4).

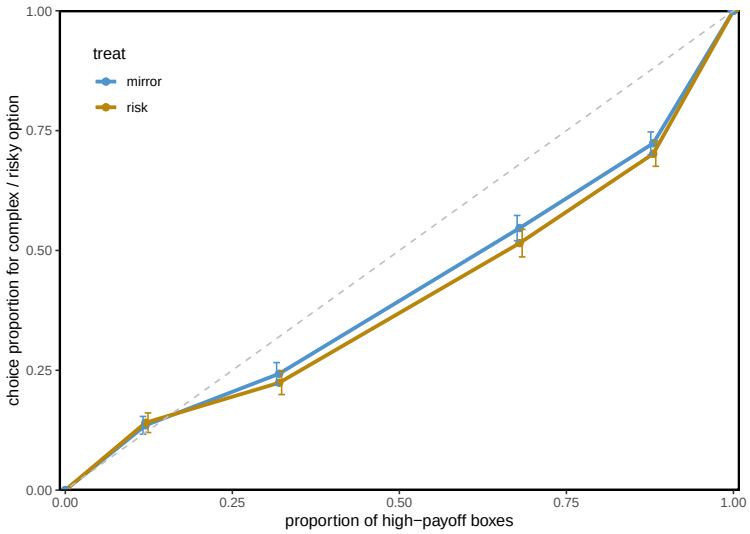


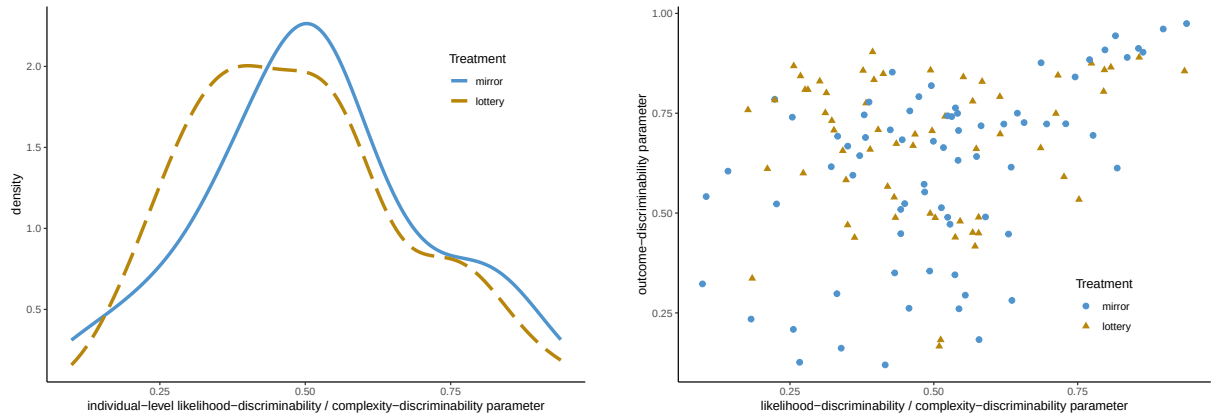
Figure 4 Choice proportions for risky or complex option

The graph plots the number of high-payoff boxes on the x-axis against choice proportions of the risky option (random box treatment) or the complex option (average box treatment). Given the equal steps in sure options, the graph can be interpreted as a (stochastic) variant of a dual-EU probability-distortion function.

Figure 4 shows the choice proportions for the risky option (random box treatment) or the complex option (average box treatment) for tasks providing different numbers of boxes associated with €18, or else 0. In the random box treatment, the data reveal an inverse-S shape as typically found in the PT literature. Strikingly, the same pattern occurs in the mirror treatment, where no risk is present and choices are between deterministic options. There is no difference in any of the choice proportions, and the same holds for choice tasks not shown in the graph (see table 1). Risk and choice complexity in a riskless setup thus produce virtually identical results.

Next, I estimate equation (11) on the complete data to examine the individual-level parameters. Figure 5, panel 5(a), shows a density plot of the individual-level likelihood-discriminability parameters (lottery treatment) and the individual-level 'complexity-discriminability' parameters (mirror treatment). Both distributions present a peak around 0.5, and look very similar. The lottery-discriminability has some probability mass to the left of the complexity-discriminability distribution, possibly indicating an effect of risk *on top* of the complexity effect, as also discussed by Oprea (2022). That effect, however, appears to be relatively small.

Panel 5(b) shows a scatter plot of the parameters capturing sensitivity towards the number of boxes, and sensitivity towards outcomes. There is no indication that behaviour for mirrors might converge towards behaviour resembling value maximization (i.e., a deterministic choice for the



(a) Density of γ for lotteries and mirrors

(b) Scatter plot of γ and α

Figure 5 Individual-level parameters for lotteries and mirrors

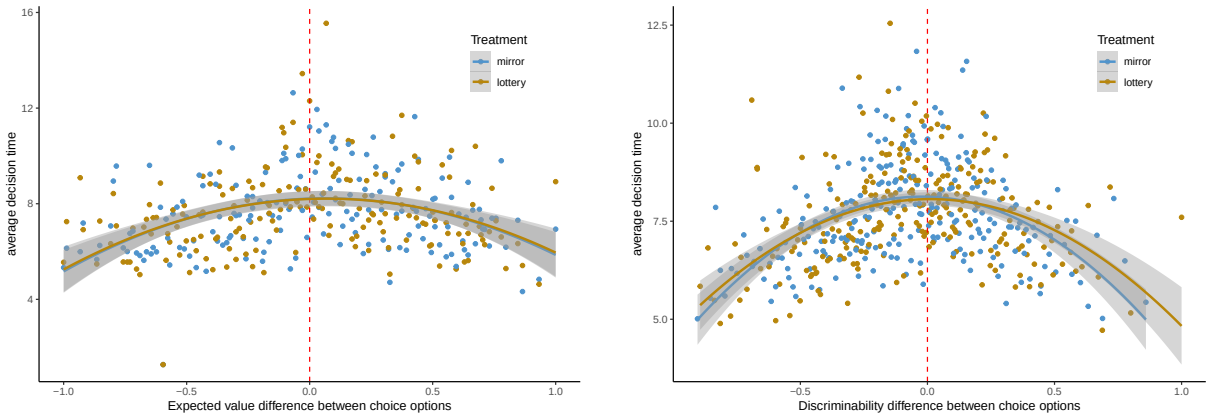
Graphical representation of individual-level parameters for risk and mirrors. The density functions in panel a have been obtained by plotting the individual-level mean estimates from the Bayesian hierarchical model. Panel b presents a scatter plot of the mean individual-level estimates of likelihood-discriminability γ and outcome-discriminability α.

option with the higher value). Remarkably, the results not only show equal distortions driven by the number of boxes across the two treatments, but also equal distortions of *outcomes*, providing direct evidence that outcome-discriminability is also driven by the complexity of the situation, rather than by any inherent attitude towards risk. This results in a clear violation of PT, and is indeed hard to reconcile with any deterministic model of risky choice. The BIM, however, can handle them easily since it models the perception of numerical quantities, rather than attitudes towards risk *per se*. While the BIM shares this feature with other models from the noisy coding family (Steiner and Stewart 2016, Netzer et al. 2021), it has an edge when it comes to quantifying patterns occurring in data, since it is geared towards direct econometric implementability.

This is further showcased by a comparison of the predictive performance of the two models on the data. The BIM will then deploy one and the same modelling setup for both treatments, whereas PT will directly compare the utilities of the two deterministic outcomes in the mirror treatment, plus an additive error term just like the one used for risk. I test model performance based on their out-of-sample predictive ability (Gelman et al. 2014a, Vehtari et al. 2017). An aggregate test yields a clear verdict in favour of the Bayesian Inference model (ELPD difference of -7653.5, with a standard error of 239.6). To take the hierarchical nature of the data into account, I also conduct a stratified *leave-k-out* test, whereby a subset of $k\%$ of the choices of each individual are removed and predicted based on the remaining $(100-k)\%$ of the observations. This procedure is repeated for each subset of $k\%$ of the data. Setting $k=10$, the results once again indicate that the BIM outperforms PT (ELPD diff. of -753.8, $se=51.9$).

Finally, let us take a look at decision times across the two treatments. Decision times are related to decision difficulty, and have been shown to be predictive of utility differences (Alós-Ferrer et al.

2021, Alós-Ferrer and Garagnani 2022). Based on standard models of decision making under risk, we would expect decision times for lotteries and mirrors to look rather different. Decision times in risky choice have been shown to be closely linked to differences in the expected utilities of choice options, whereas they tend to be only weakly linked to differences in expected values (Alós-Ferrer and Garagnani 2022). If the results were purely caused by errors in calculations of values in mirrors then we should expect to see differences between the conditions, with decision times in the mirror treatment peaking towards the point of EV equality between the two options. The fact that the distributions are very similar, however, points to a similarity in decision processes across the two settings. According to the BIM, we should not see major differences between the treatments. In both cases, the BIM thus predicts decision time to follow an inverted-U shape, with its peak at the point of no discriminability, where the two choice options become difficult to tease apart.



(a) Decision time and EV difference (b) Decision time and BIM discriminability

Figure 6 Average decision time as a function of EV difference and discriminability

Plot of expected value difference between the choice options (panel 6(a)) and normalized individivudual-level discriminability between the choice options (panel 6(b)) against average decision time. Both differences in EV and differences in discriminability have been normalized to fall between −1 and 1 at the individual level. The discriminability plot has been obtained by means of a binning procedure applied to discriminability. The solid lines indicate a least squares fit of a second-degree polynomial with surrounding uncertainty regions.

Figure 6 plots average decision times against the difference in expected values (panel 6(a)) and against the discriminability given by the nominator of (11) (panel 6(b)). Decision times increase towards the point of equality in expected values, continue to increase slightly thereafter, and subsequently decline. Importantly, there is no difference whatsoever in this trend for lotteries and mirrors, as deterministic models of decision making under risk would suggest. The inverse-U shape as a function of discriminability in the Bayesian inference model shown in panel 6(b) is much more pronounced than the curve for the EV difference. The peak in decision times coincides with the point of zero discriminability, where both options appear equally good, and shows no difference

between lotteries and mirrors. This further supports the conclusion that decisions in lotteries and mirrors are driven by the same process, in contradiction to standard models of risky choice, but in line with the predictions of the Bayesian Inference Model.

4. Discussion and Conclusion

Traditional models of decision-making under risk and uncertainty represent behaviour by deterministic preference functionals applied to objectively perceived stimuli. The deterministic choice model is then augmented by an independently chosen stochastic model to allow for preference parameters to be recovered from noisy data. The combination between decision model and stochastic model of choice, however, is arbitrary. Such arbitrary combinations of decision models and stochastic choice architectures may bias inferences on the model parameters and even about the correct underlying model if the stochastic assumptions are not warranted.

The Bayesian Inference Model I presented in this paper turns this process on its head. Starting from the premise that the choice stimuli themselves are encoded by noisy mental signals, I have shown how the optimal decoding of these signals by means of a mental prior will result in decisions that deviate systematically from the underlying optimal choice rule, and in particular, which will result in systematic outcome- and likelihood-distortions akin to those documented in the prospect theory literature. While providing cognitive micro-foundations for PT-like choice patterns, the model simultaneously predicts violations of 'PT plus white noise'. Other than PT, the model is not specifically geared at describing attitudes towards risk and uncertainty, but rather captures the noisy perception of numerical quantities. The model is thus applicable not only to lottery choice, but also to the deterministic mirrors of those lotteries introduced by Oprea (2022). At the same time, the model blurs the line between traditional categories such as risk (known probabilities) and ambiguity (unknown or vague probabilities), which according to the BIM are located along a continuum of perceptual difficulty of probabilistic information.

In the present paper, I have presented three empirical illustrations of patterns predicted by the BIM, and which are difficult to organize under traditional behavioural models of choice. The core strength of the model, however, is that it provides common foundations for decision-making phenomena which are currently explained by different behavioural models. In a twin paper (Vieider 2021), I show that an identical modelling setup as used in this paper can account for the co-existence of time discounting patterns that cannot be organized under standard models of discounting. L'Haridon et al. (2023) show that the same model can account for the Ellsberg paradox when applied to choices between risky and ambiguous wagers, and that it predicts both patterns such as predicted by prospect theory, and patterns as predicted in the multiple priors literature, thus potentially unifying these two streams of ambiguity literatures. Oprea and Vieider (2023) further

show that allowing subjects to sample from fully described choice options leads to learning, and reduces probability distortions. The occurrence of learning when probabilities are objectively given and fully described directly speaks to the perceptual uncertainty enshrined in the BIM, since learning should have no effect in the absence of uncertainty about described objective probabilities.

An important question concerns the model's status vis-a-vis prospect theory. A prime insight arising from the model is that functionals as documented in the extensive PT literature can be generated by noise in the perception of choice quantities, which need to be inferred from noisy signals. The same noise, in turn, will violate the 'PT plus white noise' assumption underlying standard empirical implementations of PT, creating systematic correlations between decision noise and PT-preference parameters estimated under that assumption. Given its ability to capture this feature, the BIM represents a promising tool for empirical investigations, particularly since it allows for the parameters of the priors and likelihoods to be endogenized in a dynamic setting. The development of such a fully endogenous model, however, will need to be left for future work.

In terms of economic theory, however, PT continues to have an edge due to its comparative simplicity. Where the parameters of the BIM are tricky to manipulate theoretically due to their inherent interdependence, the nominal independence of the equivalent PT parameters allows for a flexible modelling framework that can nevertheless capture many of the central behavioural features. At the same time, the BIM shows some of the limitations of the standard PT model, which could prove valuable in informing modelling exercises going forward.

The model proposed in this paper shares a common intuition with other models of adaptive behaviour. Most notable is the connection to the evolutionary models of Robson (2001) and Netzer (2009), who propose a fitness-maximizing model predicated on limited discernibility of outcomes to derive a utility function that adapts to the local environment. A common element is that all these models incorporate the intuition of just noticeable differences in utility, although in the BIM this is a result of the compressed mental representation of stimuli, whereas it is a modelling assumption in the Robson-Netzer framework, where it results in utility taking the form of a step function.

There are also commonalities with other models of noisy perception. Steiner and Stewart (2016) present a model in which probabilities of binary wagers are subject to noisy perception, similar to the intuition developed here. Other than in the model presented here, however, outcomes are processed without uncertainty, and the model makes no predictions about decision noise. The sure option is the quantity driving probability distortions, rather than a weight capturing the DM's confidence in the evidence. Netzer et al. (2021) present a model in which an agent receives noisy signals about different lottery arms. The agent may thereby decide to oversample some lottery arms, which could lead to the overweighting of small probabilities. While some of the underlying intuitions are similar to those developed in this model, the focus is different. Netzer et al. (2021)

primarily focus on the question of what happens when decision noise goes to zero, resulting in insights that are complementary to those presented in this paper.

The general framework based on the noisy neural coding of stimuli—and its extension to setups allowing for the adaptation of the model parameters to changes in the environment—presents the promise of a unifying theory of individual choice behaviour. On the one hand, coding noise plays a central role in the model, driving deviations from optimal behaviour, such as expected value maximization for small-stake risks. This results in behavioural regularities that are likely to impact decisions far beyond the particular setup used in this paper. For instance, we may well expect the presentation format of stimuli to impact choices, and the noise of stimulus encoding may be expected to increase systematically with the difficulty of the choice tasks. The precise implications of this insight deserve close attention, and are thus left for future work.

Possibly the most far-reaching implications, however, concern the role of the prior mean. A dynamic hierarchical model detailing how the prior is determined promises to shed light on how preferences adapt to new choice situations, as documented empirically by Di Falco and Vieider (2022). Understanding both short term determinants of the prior in the lab, and long-term determinants of the prior in the real world should thus be a priority topic for future research. Notice that particularly long-term components of the prior—likely contained in a higher-level prior that can inform decisions and behaviour across a variety of choice situations—could present a unique chance to deliver a coherent theory of preference formation.

Acknowledgments

I gratefully acknowledge financial support from the Research Foundation—Flanders (FWO) under the project "Causal Determinants of Preferences" (G008021N), and the special research fund (BOF) at Ghent University under the project "The role of noise in the determination of risk preferences". A previous version of this paper has been circulated under the title "Noisy Coding and Decisions under Uncertainty". I am indebted to Mohammed Abdellaoui, Aurélien Baillon, Ranoua Bouchouicha, Benjamin Enke,Thomas Epper, Ernst Fehr, Cary Frydman, Olivier L'Haridon, Paulo Natenzon, Nick Netzer, Ryan Oprea, Pietro Ortoleva, Jakub Steiner, Tom Verguts, Peter Wakker, and Horst Zank for helpful comments and discussions. All errors remain my own.

Appendix A: Derivation for multi-outcome wagers

I assume that each term in the comparison in (4) is logged and evaluated in parallel within a neural network. That implies that there will be a comparison between log-odds and log-cost benefits just like the one discussed in the main text for each comparative state of the world. For an arbitrary comparison i, we will thus observe the following likelihood functions:

$$r_e^i \sim \mathcal{N}\left(ln\left(\frac{P[e_i]}{P[e_n]}\right), \nu^2\right) \quad , \quad r_o^i \sim \mathcal{N}\left(\mathbb{1}\, ln\left(\frac{\mathbb{1}(x_i - y_i)}{y_i - x_n}\right), \nu^2\right). \tag{15}$$

where $\mathbb{1}=1$ if $x_i > y_i$, and $\mathbb{1}=-1$ if $x_i < y_i$ ($x_n < y_i \ \forall i$ by construction of the comparison). Other than in the binary case, the signal r_o should now be understood as the signal for the *relative benefit-cost ratio*, since one can no longer simply refer to 'costs' and 'benefits' in the general case.

We can follow the same steps as in the binary case to obtain a threshold equation and a contribution to the overall evidence in favour of taking the riskier wager from each of the single comparisons. This will result in the following multidimensional generalization of (11):

$$P[\boldsymbol{x} \succ \boldsymbol{y}] = \Phi\left[\sum_{i=1}^{n-1} \frac{\gamma \times ln\left(\frac{P[e_i]}{P[e_n]}\right) + \mathbb{1}\,\alpha \times ln\left(\frac{\mathbb{1}(x_i - y_i)}{y_i - x_n}\right) - ln(\delta)^{-1}}{\nu \times \sqrt{(n-1)\times(\alpha^2+\gamma^2)}}\right]. \tag{16}$$

The result obtains from the observation that, since the single contributions to the comparative evaluation of the wagers follow a standard normal distribution, their sum will also follow a standard normal distribution. Note that this decision rule is thus fully general and applicable to all forced choice paradigms, as long as n is finite (as will typically be the case for monetary outcomes specified down to the last cent). Other than in modern, cumulative, version of PT, the probability distortions are not applied to good-news probabilities attached to ranked outcomes, so that for more than 2 non-zero comparative outcomes the model provides microfoundations for the biseparable model of Ghirardato et al. (2004).

References

Abdellaoui, Mohammed, Aurélien Baillon, Lætitia Placido, and Peter P. Wakker (2011) 'The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation.' *American Economic Review* 101, 695–723

Agranov, Marina, and Pietro Ortoleva (2017) 'Stochastic choice and preferences for randomization.' *Journal of Political Economy* 125(1), 40–68

Alós-Ferrer, Carlos, and Michele Garagnani (2022) 'Strength of preference and decisions under risk.' *Journal of Risk and Uncertainty* pp. 1–21

Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer (2021) 'Time will tell: Recovering preferences when choices are noisy.' *Journal of Political Economy* 129(6), 1828–1877

Apesteguia, Jose, and Miguel A Ballester (2018) 'Monotone stochastic choice models: The case of risk and time preferences.' *Journal of Political Economy* 126(1), 74–106

Baillon, Aurélien, Han Bleichrodt, Umut Keskin, Olivier L'Haridon, and Chen Li (2018a) 'The effect of learning on ambiguity attitudes.' *Management Science* 64(5), 2181–2198

Baillon, Aurélien, Zhenxing Huang, Asli Selim, and Peter P Wakker (2018b) 'Measuring ambiguity attitudes for all (natural) events.' *Econometrica* 85, 1839–1858

Bhui, Rahul, Lucy Lai, and Samuel J Gershman (2021) 'Resource-rational decision making.' *Current Opinion in Behavioral Sciences* 41, 15–21

Blavatskyy, Pavlo R (2014) 'Stronger utility.' *Theory and Decision* 76(2), 265–286

Bossaerts, Peter, Nitin Yadav, and Carsten Murawski (2019) 'Uncertainty and computational complexity.' *Philosophical Transactions of the Royal Society B* 374(1766), 20180138

Bouchouicha, Ranoua, and Ferdinand M. Vieider (2017) 'Accommodating stake effects under prospect theory.' *Journal of Risk and Uncertainty* 55(1), 1–28

Bouchouicha, Ranoua, Jilong Wu, and Ferdinand M. Vieider (2023) 'Choice lists and 'standard patterns' of risk taking.' Technical Report, Ghent University Discussion Papers

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper (2010) 'Risk and Rationality: Uncovering Heterogeneity in Probability Distortion.' *Econometrica* 78(4), 1375–1412

Buschena, David, and David Zilberman (2000) 'Generalized expected utility, heteroscedastic error, and path dependence in risky choice.' *Journal of Risk and Uncertainty* 20(1), 67–88

Busemeyer, Jerome R, and James T Townsend (1993) 'Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment.' *Psychological review* 100(3), 432

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell (2017) 'Stan: A probabilistic programming language.' *Journal of Statistical Software* 76(1), 1–32

Choi, Syngjoo, Jeongbin Kim, Eungik Lee, Jungmin Lee et al. (2022) 'Probability weighting and cognitive ability.' *Management Science, forthcoming* 68(7), 4755–5555

Chuang, Yating, and Laura Schechter (2015) 'Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results.' *Journal of Development Economics* 117, 151–170

Dayan, Peter, and Laurence F Abbott (2001) *Theoretical neuroscience: computational and mathematical modeling of neural systems* (Computational Neuroscience Series)

Dehaene, Stanislas (2003) 'The neural basis of the weber–fechner law: a logarithmic mental number line.' *Trends in cognitive sciences* 7(4), 145–147

Dehaene, Stanislas, and Jean-Pierre Changeux (1993) 'Development of elementary numerical abilities: A neuronal model.' *Journal of cognitive neuroscience* 5(4), 390–407

Di Falco, Salvatore, and Ferdinand M. Vieider (2022) 'Environmental adaptation of risk preferences.' *Economic Journal* 132(648), 2737–2766

Dimmock, Stephen G., Roy Kouwenberg, and Peter P. Wakker (2015) 'Ambiguity Attitudes in a Large Representative Sample.' *Management Science* 62(5), 1363–1380

Doya, Kenji, Shin Ishii, Alexandre Pouget, and Rajesh PN Rao (2007) *Bayesian brain: Probabilistic approaches to neural coding* (MIT press)

Einstein, Albert (1922) *Sidelights on relativity* (Methuen & Company Limited)

Ellsberg, Daniel (1961) 'Risk, Ambiguity and the Savage Axioms.' *Quarterly Journal of Economics* 75(4), 643–669

Enke, Benjamin, and Thomas Graeber (2023) 'Cognitive uncertainty.' *Quarterly Journal of Economics*

Fechner, Gustav Theodor (1860) (Kessinger's Legacy Reprints)

Fehr-Duda, Helga, Adrian Bruhin, Thomas F. Epper, and Renate Schubert (2010) 'Rationality on the Rise: Why Relative Risk Aversion Increases with Stake Size.' *Journal of Risk and Uncertainty* 40(2), 147–180

Fehr-Duda, Helga, and Thomas Epper (2012) 'Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences.' *Annual Review of Economics* 4(1), 567–593

Gabaix, Xavier, and David Laibson (2017) 'Myopia and discounting.' Technical Report, National bureau of economic research

Gelman, Andrew, and Jennifer Hill (2006) *Data analysis using regression and multilevel/hierarchical models* (Cambridge university press)

Gelman, Andrew, Jessica Hwang, and Aki Vehtari (2014a) 'Understanding predictive information criteria for bayesian models.' *Statistics and computing* 24(6), 997–1016

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin (2014b) *Bayesian data analysis,* vol. 2 (CRC press Boca Raton, FL)

Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci (2004) 'Differentiating ambiguity and ambiguity attitude.' *Journal of Economic Theory* 118(2), 133–173

Gold, Joshua I, and Michael N Shadlen (2001) 'Neural computations that underlie decisions about sensory stimuli.' *Trends in cognitive sciences* 5(1), 10–16

Goldstein, W, and H Einhorn (1987) 'Expression Theory and the Preference Reversal Phenomena.' *Psychological Review* 94, 236–254

Gonzalez, Richard, and George Wu (1999) 'On the Shape of the Probability Weighting Function.' *Cognitive Psychology* 38, 129–166

Green, David Marvin, John A Swets et al. (1966) *Signal detection theory and psychophysics,* vol. 1 (Wiley New York)

Gul, Faruk, and Wolfgang Pesendorfer (2006) 'Random expected utility.' *Econometrica* 74(1), 121–146

Harvey, Ben M, Barrie P Klein, Natalia Petridou, and Serge O Dumoulin (2013) 'Topographic representation of numerosity in the human parietal cortex.' *Science* 341(6150), 1123–1126

Heath, C, and A Tversky (1991) 'Preference and belief: Ambiguity and competence in choice under uncertainty. Journal of Risk and.' *Journal of Risk and Uncertainty* 4(1), 5–28

Herold, Florian, and Nick Netzer (2023) 'Second-best probability weighting.' *Games and Economic Behavior* 138, 112–125

Hey, John D., and Chris Orme (1994) 'Investigating Generalizations of Expected Utility Theory Using Experimental Data.' *Econometrica* 62(6), 1291–1326

Hogarth, Robin M., and Hillel J. Einhorn (1990) 'Venture Theory: A Model of Decision Weights.' *Management Science* 36(7), 780–803

Holt, Charles A., and Susan K. Laury (2002) 'Risk Aversion and Incentive Effects.' *American Economic Review* 92(5), 1644–1655

Howard, Marc W, and Karthik H Shankar (2018) 'Neural scaling laws for an uncertain world.' *Psychological review* 125(1), 47

Kahneman, Daniel, and Amos Tversky (1979) 'Prospect Theory: An Analysis of Decision under Risk.' *Econometrica* 47(2), 263 – 291

Khaw, Mel Win, Ziang Li, and Michael Woodford (2021) 'Cognitive imprecision and small-stakes risk aversion.' *The Review of Economic Studies* 88(4), 1979–2013

Knill, David C, and Alexandre Pouget (2004) 'The bayesian brain: the role of uncertainty in neural coding and computation.' *TRENDS in Neurosciences* 27(12), 712–719

Körding, Konrad Paul, and Daniel M Wolpert (2004) 'The loss function of sensorimotor learning.' *Proceedings of the National Academy of Sciences* 101(26), 9839–9842

Kruschke, John (2014) *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (Academic Press)

L'Haridon, Olivier, and Ferdinand M. Vieider (2019) 'All over the map: A worldwide comparison of risk preferences.' *Quantitative Economics* 10, 185–215

L'Haridon, Olivier, Ferdinand M. Vieider, Diego Aycinena, Agustinus Bandur, Alexis Belianin, Lubomir Cingl, Amit Kothiyal, and Peter Martinsson (2018) 'Off the charts: Massive unexplained heterogeneity in a global study of ambiguity attitudes.' *Review of Economics and Statistics* 100(4), 664–677

L'Haridon, Olivier, Ryan Oprea, Rafael Polania, and Ferdinand M. Vieider (2023) 'Cognitive foundations of ambiguity atttitudes.' *Mimeo*

Loomes, Graham (2010) 'Modeling choice and valuation in decision experiments.' *Psychological review* 117(3), 902

Ma, Wei Ji, Konrad Paul Kording, and Daniel Goldreich (2023) *Bayesian Models of Perception and Action: An Introduction* (MIT press)

McElreath, Richard (2016) *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (Academic Press)

Mosteller, Frederick, and Philip Nogee (1951) 'An experimental measurement of utility.' *Journal of Political Economy* 59(5), 371–404

Natenzon, Paulo (2019) 'Random choice and learning.' *Journal of Political Economy* 127(1), 419–457

Netzer, Nick (2009) 'Evolution of time preferences and attitudes toward risk.' *American Economic Review* 99(3), 937–55

Netzer, Nick, Arthur Robson, Jakub Steiner, and Pavel Kocourek (2021) 'Endogenous risk attitudes.' Working Paper

Nieder, Andreas, and Earl K Miller (2003) 'Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex.' *Neuron* 37(1), 149–157

Oprea, Ryan (2022) 'Simplicity equivalents.' *Working Paper*

Oprea, Ryan, and Ferdinand M. Vieider (2023) 'Closing the gap.' *Mimeo*

Petzschner, Frederike H, Stefan Glasauer, and Klaas E Stephan (2015) 'A bayesian perspective on magnitude estimation.' *Trends in cognitive sciences* 19(5), 285–293

Piazza, Manuela, Véronique Izard, Philippe Pinel, Denis Le Bihan, and Stanislas Dehaene (2004) 'Tuning curves for approximate numerosity in the human intraparietal sulcus.' *Neuron* 44(3), 547–555

Robson, Arthur J (2001) 'The biological basis of economic behavior.' *Journal of Economic Literature* 39(1), 11–33

Savage, Leonard J. (1954) *The Foundations of Statistics* (New York: Wiley)

Stan Development Team (2017) 'RStan: the R interface to Stan.' R package version 2.17.2

Steiner, Jakub, and Colin Stewart (2016) 'Perceiving prospects properly.' *American Economic Review* 106(7), 1601–31

Trautmann, Stefan T., and Gijs van de Kuilen (2015) 'Ambiguity Attitudes.' In 'The Wiley Blackwell Handbook of Judgment and Decision Making' (Wiley Blackwell)

Tversky, Amos (1969) 'Intransitivity of preferences.' *Psychological review* 76(1), 31

Tversky, Amos, and Daniel Kahneman (1992) 'Advances in Prospect Theory: Cumulative Representation of Uncertainty.' *Journal of Risk and Uncertainty* 5, 297–323

Vehtari, Aki, Andrew Gelman, and Jonah Gabry (2017) 'Practical bayesian model evaluation using leave-one-out cross-validation and waic.' *Statistics and computing* 27(5), 1413–1432

Vieider, Ferdinand M (2018) 'Violence and risk preference: Experimental evidence from afghanistan: Comment.' *American Economic Review* 108(8), 2366–82

Vieider, Ferdinand M. (2021) 'Noisy coding of time delays and reward discounting.' Working Paper21/1036 FEB, Ghent University, Belgium

Vilares, Iris, and Konrad Kording (2011) 'Bayesian models: the structure of the world, uncertainty, behavior, and the brain.' *Annals of the New York Academy of Sciences* 1224(1), 22

Wakker, Peter P. (2010) *Prospect Theory for Risk and Ambiguity* (Cambridge: Cambridge University Press)

Wilcox, Nathaniel T. (2011) ''Stochastically more risk averse:' A contextual theory of stochastic discrete choice under risk.' *Journal of Econometrics* 162(1), 89–104

Woodford, Michael (2012) 'Prospect theory as efficient perceptual distortion.' *American Economic Review* 102(3), 41–46

Yaari, Menahem E. (1987) 'The Dual Theory of Choice under Risk.' *Econometrica* 55(1), 95–115

Zeisberger, Stefan, Dennis Vrecko, and Thomas Langer (2012) 'Measuring the time stability of Prospect Theory preferences.' *Theory and Decision* 72(3), 359–386

Zhang, Hang, and Laurence T Maloney (2012) 'Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition.' *Frontiers in neuroscience* 6, 1

ONLINE APPENDIX
Decisions under Uncertainty as Bayesian Inference

Appendix S1: Derivation for unlogged choice rule

The derivations in the main text as well as the details in the last subsection were based on the logarithmic choice rule in equation 3. If we were to use the median of the posterior instead of its mean in the choice rule, then the derivations would be identical if we were to base them on the unlogged choice rule in equation 1 instead. Körding and Wolpert (2004) show that this is indeed optimal if the absolute value of the error is used as a loss function. Given that the derivation in the main text is based on the mean, however, the two choice rules will result in a slightly different interpretation of the prior mean. I here present the derivation based on the unlogged choice rule. We then have the following mental choice rule:

$$E\left[\frac{P[e]}{P[\tilde{e}]}|r_e\right] > E\left[\frac{c-y}{x-c}|r_o\right]. \tag{17}$$

If we want to use the posterior expectation in the choice rule instead of the median, as done here, then we need to obtain the posterior means of the unlogged quantities by taking the exponential of the expectations in equation 7 in the main text:

$$E\left[\left(\frac{P[e]}{P[\tilde{e}]}\right)|r_e\right] = e^{\gamma\times r_e + ln(\xi) + \frac{1}{2}\widehat{\sigma}_e^2} \;\;,\;\; E\left[\left(\frac{c-y}{x-c}\right)|r_o\right] = e^{\alpha\times r_o + ln(\zeta) + \frac{1}{2}\widehat{\sigma}_o^2}, \tag{18}$$

where $\widehat{\sigma}_e^2 \triangleq \frac{\nu^2\sigma_e^2}{\nu^2+\sigma_e^2}$ and $\widehat{\sigma}_o^2 \triangleq \frac{\nu^2\sigma_o^2}{\nu^2+\sigma_o^2}$ are the posterior variances. We can now substitute these quantities into the mental choice rule above, to obtain:

$$e^{\gamma\times r_e + ln(\xi) + \frac{1}{2}\widehat{\sigma}_e^2} > e^{\alpha\times r_o + ln(\zeta) + \frac{1}{2}\widehat{\sigma}_o^2} \tag{19}$$

Taking the logarithm of both sides, defining $\widehat{\xi} \triangleq e^{ln(\xi)+\frac{1}{2}\widehat{\sigma}_e^2}$ and $\widehat{\zeta} \triangleq e^{ln(\zeta)+\frac{1}{2}\widehat{\sigma}_o^2}$, we obtain:

$$\gamma\times r_e - \alpha\times r_o > ln(\widehat{\delta})^{-1}, \tag{20}$$

where $\widehat{\delta} \triangleq \widehat{\xi}\times\widehat{\zeta}^{-1}$. All further derivations proceed like above. The sole difference with the results presented in the main text thus flow from the difference in definition of $\widehat{\delta}$ versus δ. That is, the variance of the priors will enter into the definition of $\widehat{\delta}$, while it will not enter the definition of δ. In the main text, I argued that the logged choice rule appears more plausible than its unlogged version for computational reasons. The less intuitive formulation of the prior mean emerging from the unlogged choice rule may constitute a further argument for my preferred setup. That being said, the difference between the two setups is *quantitative* in nature, rather than *qualitative*, and it does not affect the main conclusions drawn in the paper.

Appendix S2: Additional results on PT correlations

This section presents additional details and results for the analysis of 2-outcome wagers under risk in a PT setting.

S2.1. Estimation of the Bayesian hierarchical model

I estimate individual-level parameters using Bayesian hierarchical models (Gelman and Hill 2006, Gelman et al. 2014b, McElreath 2016). I conduct the analysis using Stan (Carpenter et al. 2017). Take a parameter vector $\boldsymbol{\theta}_i$, indicating individual-level parameters. This vector follows the following distribution:

$$\boldsymbol{\theta}_i \sim \mathcal{N}(\overline{\boldsymbol{\theta}}, \Sigma), \tag{21}$$

where $\overline{\boldsymbol{\theta}}$ is a vector of hyperparameters containing the means of the individual-level parameters, and Σ is a variance-covariance matrix of the individual-level parameters.[13] I estimate the parameters in Stan launched from Rstan (Stan Development Team 2017). Estimations typically employ 4 chains with 2000 iterations per chain, of which 1000 are warmup iterations—the default settings of Stan. I checked convergence by examining the R-hat statistics, and by checking for divergent iterations. The model endogenously estimates the priors for the individual-level parameters from the aggregate data, resulting in partial pooling. This is indeed a central strength of the model, which tends to reduce issue with overfitting of few observations. The priors for the estimation of the aggregate-level means are chosen in such a way as to be mildly regularizing (McElreath 2016). That is, their variance is chosen in such a way that all plausible parameters fall into a region attributed high likelihood, but narrow enough to nudge the simulation algorithm towards convergence. In any case, the datasets I use have sufficient data at the aggregate level for the priors chosen to have little or no impact on the final result.

Following the estimation approach used in Bruhin et al. (2010), I use the density around the observed switching point to estimate the model. The model takes the following form:

$$ce \sim \mathcal{N}(u^{-1}[w(p)u(x) + (1 - w(p))u(y)], \widehat{\omega}^2 \times |x - y|), \tag{22}$$

with $u(x) = x^{\widehat{\alpha}}$ and $w(p) = \frac{\widehat{\delta} p^{\widehat{\gamma}}}{\widehat{\delta} p^{\widehat{\gamma}} + (1-p)^{\widehat{\gamma}}}$, and where multiplying the variance $\widehat{\omega}$ by $|x - y|$ allows for heteroscedasticity across choice lists with different step sizes between choices, following the approach in the original paper from which I took the data. Contextualizing choices by letting the error be heteroscedastic across to the *utility difference*, $|u(x) - u(y)|$, such as proposed by Wilcox (2011), does not affect the conclusions I draw.

The priors are chosen such as to be informative about the expected location of the model parameters, without imposing any undue restrictions on the data. This is typically referred to as *mildly regularizing* priors, and it helps convergence in the model. For instance, the prior chosen for $\widehat{\gamma}$ has a mean of 0.7 on the original scale, with 95% of the probability mass allocated to a range of $[0, 3.92]$. This can be expected to encompass most likely parameter values. Given furthermore the large quantity of data present at the aggregate level, the data can easily overpower the prior even for parameters falling outside this range. Making the prior more diffuse and shifting the mean to e.g. 1, does not affected the estimated parameters in any way, showing that the prior has only a minimal influence on the ultimate parameter estimates. The full model in Stan is as follows (@to be added upon publication):

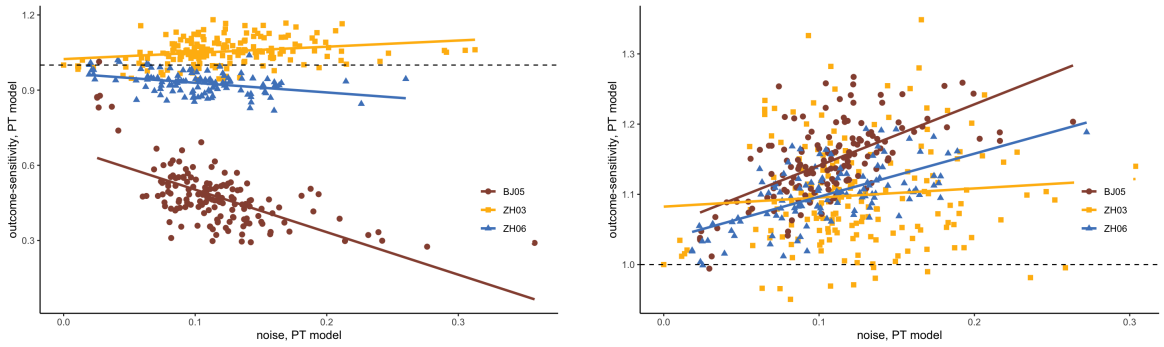
[13] Note that this specification did not yield meaningful estimates of the covariance parameters in the Zurich 2003 data of Bruhin et al. (2010). I thus estimated the PT model on that dataset with a variance matrix containing zeros in the off-diagonal elements.

S2.2. PT parameter correlations: Additional results

This section adds some further results to the correlations amongst PT parameters presented in the main text. I will present additional results for both Bruhin et al. (2010), as well as presenting equivalent results for the data of L'Haridon and Vieider (2019), for losses as well as for gains.

Figure S7 shows the correlations between the noise parameter and the outcome sensitivity parameter, $\widehat{\alpha}$, with panel 7(a) showing the results for gains. Note that in this case the BIM does not make any clear predictions, given that utility in PT is defined over individual (absolute) outcomes, whereas the BIM incorporates outcome-discriminability defined over the log-cost benefits. The correlations are thus purely shown to further document the violation of the white noise assumption.

There is a negative correlation in the Beijing 05 data ($\rho = -0.566, p < 0.001$), as well as in the Zurich 06 data ($\rho = -0.357, p < 0.001$), but a positive correlation in the Zurich 03 data ($\rho = 0.279, p < 0.001$). Noise can thus be correlated with either excess outcome sensitivity or with insensitivity towards outcomes under PT. This intuition is further confirmed for losses, shown in panel 7(b). Here we witness a positive correlation in the aggregate data ($\rho = 0.258, p < 0.001$), as well as in the three individual experiments (ZH03: $\rho = 0.07, p = 0.38$; BJ05: $\rho = 0.612, p < 0.001$; ZH06: $\rho = 0.558, p < 0.001$). This is driven by concave utility for losses in all three experiments.[14]



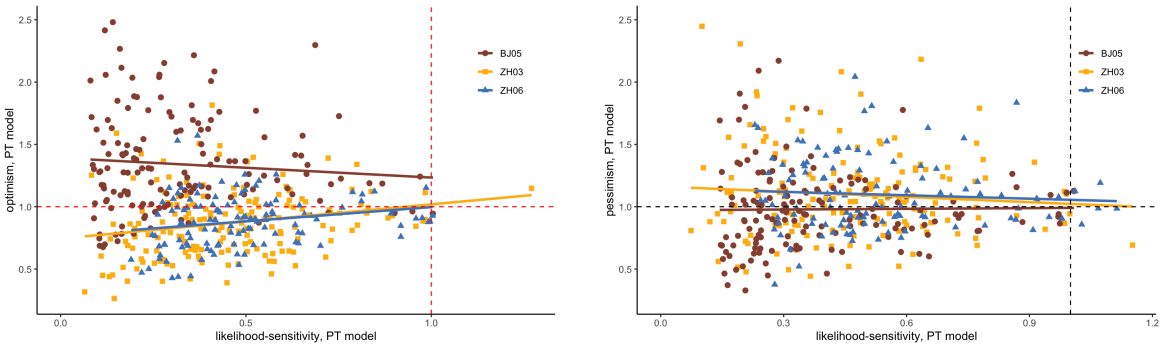
(a) Scatter plot of $\widehat{\omega}^{+}$ and $\widehat{\alpha}^{+}$ (b) Scatter plot of $\widehat{\omega}^{-}$ and $\widehat{\alpha}^{-}$

Figure S7 Scatter plot of PT parameters, outcome-sensitivity

The parameters have been obtained from the estimation of a PT model plus additive noise. The different colours and shapes represent the 3 experiments in Bruhin et al. (2010): ZH03 stands for Zurich 03; ZH6 for Zurich 06; and BJ05 for Beijing 05. The dashed lines indicate the reference parameter values of 1 (linear utility). Some outliers may be cut for better visual display.

I start by presenting correlations between additional PT parameters in the Bruhin et al. (2010) data. Another interesting pattern emerges from the relation between likelihood-sensitivity and optimism for gains, shown in figure S8 panel 8(a), and between likelihood-sensitivity and pessimism for losses, shown in panel 8(b). In both cases, the values of $\widehat{\delta}$ are most dispersed for small values of $\widehat{\gamma}$, with the dispersion decreasing

[14] Note that the BIM makes no clear predictions of correlations between noise and utility curvature in a PT setting, given the rather different definition of outcome-discriminability in the BIM. The correlations are shown in appendix S2.2 purely in order to test the 'white noise'.

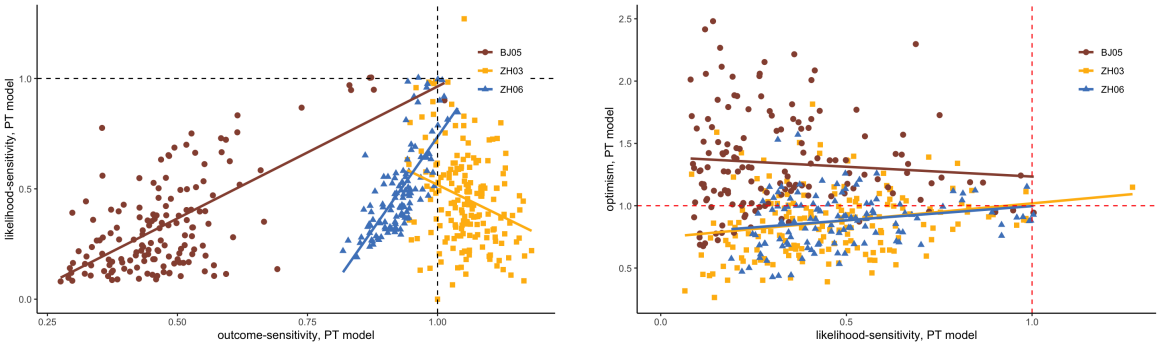


(a) Scatter plot of $\widehat{\gamma}^{+}$ and $\widehat{\delta}^{+}$ (b) Scatter plot of $\widehat{\gamma}^{-}$ and $\widehat{\delta}^{-}$

Figure S8 Scatter plot of PT parameters $\widehat{\gamma}$ and $\widehat{\delta}$

The parameters have been obtained from the estimation of a PT model plus additive noise. The different colours and shapes represent the 3 experiments in Bruhin et al. (2010): ZH03 stands for Zurich 03; ZH6 for Zurich 06; and BJ05 for Beijing 05.

markedly as $\widehat{\gamma}$ increases, resulting in a funnel with the narrow part pointing to the right. Testing the correlation of absolute deviations of $\widehat{\delta}$ from 1 with $\widehat{\gamma}$, I find highly significant effects for both gains ($\rho = -0.284, p < 0.001$) and losses ($\rho = -0.264, p < 0.001$). These patterns have no obvious explanation under PT. In the BIM, however, they are predicted by the definition of $\delta = \left(\frac{\psi}{1-\psi}\right)^{1-\gamma}$. That is, larger values of γ shift the attention from the prior to the likelihood, thus compressing δ towards 1. Panel 9(a) visualizes the correlations between $\widehat{\alpha}$ and $\widehat{\gamma}$ already discussed in the main text.



(a) Scatter plot of $\widehat{\alpha}$ against $\widehat{\gamma}$ (b) Scatter plot of $\widehat{\gamma}$ against $\widehat{\delta}$

Figure S9 Correlations between PT parameters in Bruhin et al. (2010)

Panel 9(b) further shows correlational patterns between $\widehat{\gamma}$ and $\widehat{\delta}$. According to the BIM, the results should depend on the initial value of ψ. In particular, the larger the value of $\widehat{\gamma}$, the closer the value of $\widehat{\delta}$ should be compressed towards 1. This is exactly what we observe. In experiments where we observe preponderantly values of $\widehat{\delta} < 1$, the distance to 1 decreases as $\widehat{\gamma}$ increases, thus resulting in a positive correlation between the two parameters. This is the case in the Zurich 03 experiment ($\rho = 0.262, p < 0.001$), as well as in the Zurich 06 experiment ($\rho = 0.369, p < 0.001$). In the Beijing 05 experiment, on the other hand, we observe very large

values of $\widehat{\delta} > 1$. We may thus expect a negative relationship with $\widehat{\gamma}$. We fail to observe such a relationship in the data ($\rho = 0.045, p = 0.54$). This may be due to the fact that we have very few observations with large likelihood-sensitivity in that experiment.

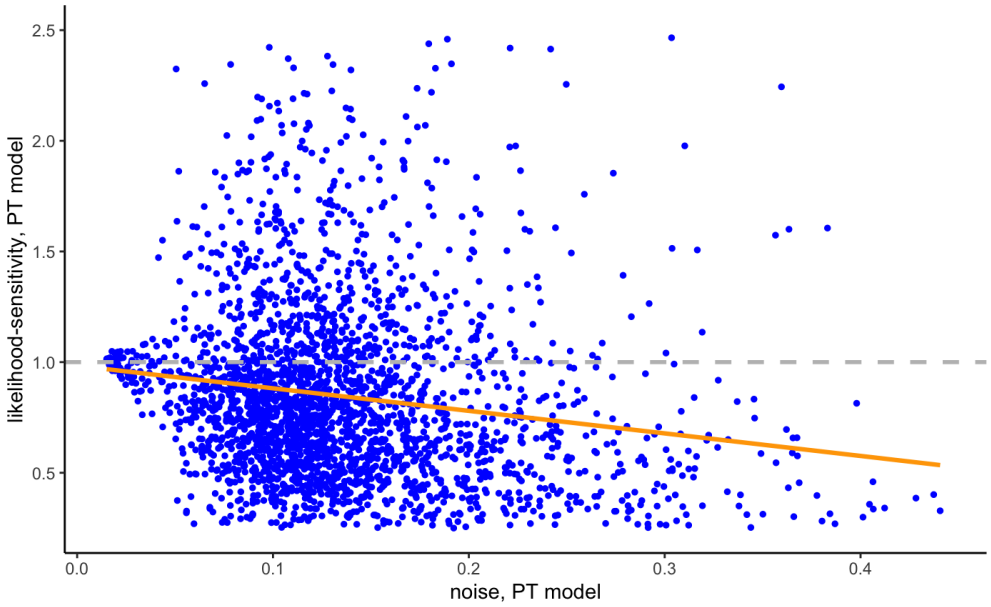


Figure S10 Correlatiion of $\widehat{\omega}^+$ and $\widehat{\gamma}^+$ in L'Haridon and Vieider (2019)

I next document correlations amongst the PT parameters in the global data of L'Haridon and Vieider (2019), for all 30 countries and 3000 subjects. Correlations are tested on parameters demeaned at the country level, corresponding to a fixed effects specification. Figure S10 shows the correlation between noise and likelihood-sensitivity for gains under PT. We find the usual negative correlation already described for the risk and rationality data ($\rho = -0.2, p < 0.001$). At first, the negative correlation may appear somewhat weaker than witnessed for Bruhin et al. (2010). Closer examination reveals that this is due to the presence of a larger number of estimates of $\widehat{\gamma}^+ > 1$. As already described for utility curvature in the main text, both negative and positive deviations tend to be correlated with noise for parameters estimated in a PT context. Taking absolute deviations of likelihood sensitivity from 1, $|1 - \widehat{\gamma}^+|$, the correlation becomes indeed much stronger ($\rho = 0.386, p < 0.001$).

The patterns for losses are very similar, and are shown in figure S11. Once again, we observe a negative correlation in the global data ($\rho = -0.167, p < 0.001$), albeit one that is not as strong as we might have expected based on the Bruhin et al. (2010) results. This is once again driven by the presence of values of $\widehat{\gamma}^- > 1$, which are also associated with high noise levels. Looking at the correlations between noise and absolute deviations of the variable from 1, $|1 - \widehat{\gamma}^-|$, the correlation thus appears much stronger ($\rho = 0.445, p < 0.001$).

Figure S12 shows the global correlation between noise and outcome sensitivity for gains. The correlation in the global data is positive, reflecting the fact that the median outcome sensitivity parameter is positive at 1.125. The correlation is highly significant ($\rho = 0.141, p < 0.001$). Results for losses are shown in figure S13, where the same patterns appear to be even more accentuated ($\rho = 0.34, p < 0.001$).

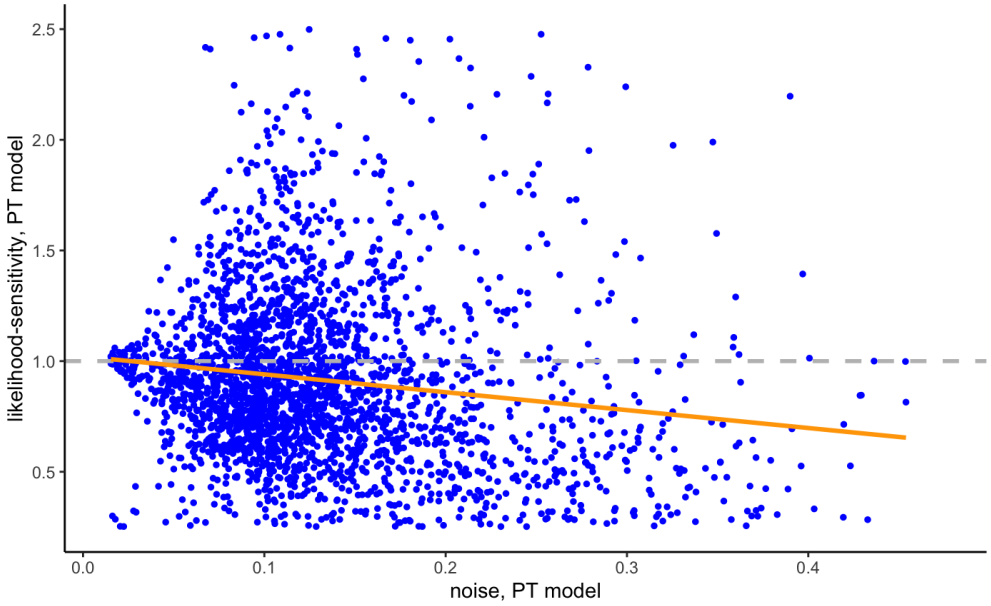


Figure S11 **Correlatiion of $\widehat{\omega}^{-}$ and $\widehat{\gamma}^{-}$ in L'Haridon and Vieider (2019)**

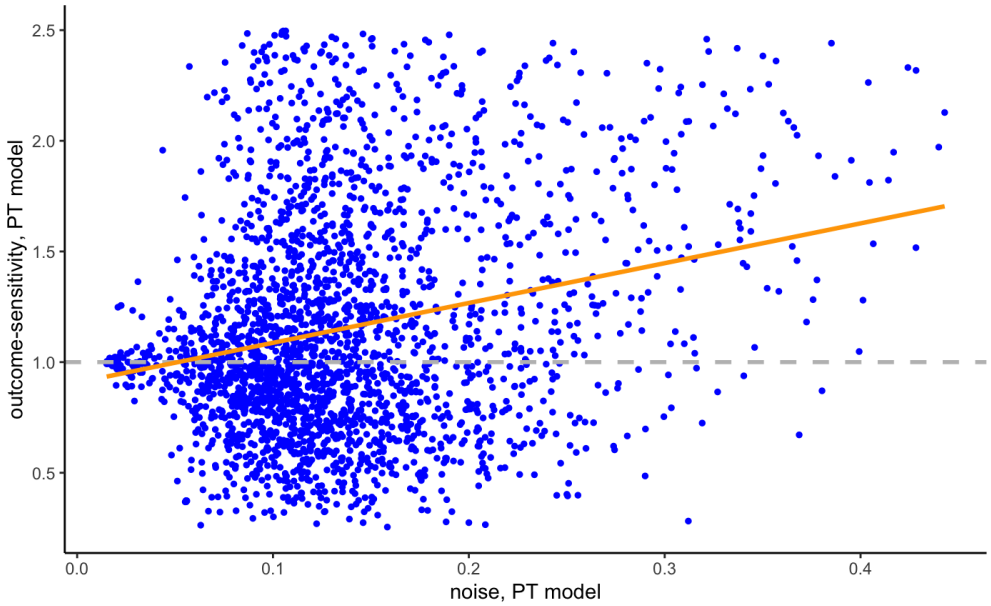


Figure S12 **Correlatiion of $\widehat{\omega}^{+}$ and $\widehat{\alpha}^{+}$ in L'Haridon and Vieider (2019)**

S2.3. The BIM beats PT in predictive accuracy

A natural next question concerns the predictive performance of the two models. This question, however, comes with a caveat. The BIM is naturally developed for binary choices. Given the additional contextual information provided by choice lists, these would need to be modelled explicitly (see Bouchouicha et al. 2023, who document large differences in behaviour arising from the choice list context). Furthermore, to be

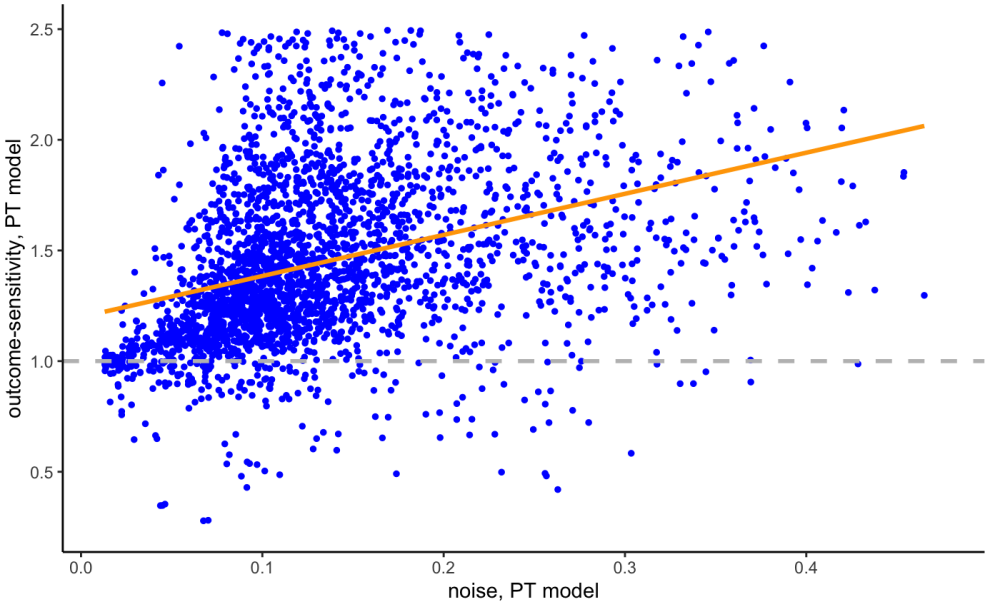


Figure S13 Correlatiion of $\widehat{\omega}^+$ and $\widehat{\alpha}^+$ in L'Haridon and Vieider (2019)

applicable to the certainty equivalence data used in the main text, the BIM needs to be rescaled, an operation that implicitly also rescales the noise. That is, the model is now estimate on equation (13), which captures the spirit of the BIM, but moves away from its literal implementation. While such rescaling is routine in PT, where the error term is arbitrary, it stretches the limits of applicability of the BIM.

All that being said, table S2 reports test statistics for Bruhin et al. (2010), separately for their three experiments, and further separated for gains and losses. In addition to the additive noise model described above, the table also presents a test of the BMI against PT including a contextual noise term, whereby the error is made proportional to the utility difference of the prospect (Wilcox 2011). The BIM outperforms PT in all 6 test cases for both additive noise and contextual noise, and the difference in performance is large. While there are yet other error specifications one could add to the PT model, this highlights precisely the arbitrariness in composition of decision model and error model discussed in the introduction, and is thus besides the point. The good performance of the BIM using the Bruhin et al. (2010) data is in no way exceptional.

Table S3 provides tests of the predictive ability of the BIM versus PT using the data of L'Haridon and Vieider (2019). I conduct the tests country by country, and separately for gains and losses. Conditional on the design, this gives me 60 independent tests. Once again, the BIM easily outperforms PT in all 60 cases, with all test statistics two orders of magnitude larger than the associated standard errors.

The BIM *predicts* the type of correlations shown above and in the main text, since the choice parameters in the numerator and the noise parameters in the denominator of equation (9) contain some of the same quantities. This is again best explained in the context of causal inference. Whereas in PT the choice characteristics in interaction with pre-existing preferences are supposed to cause behaviour (see above), in the BIM *coding noise* is the ultimate cause of behaviour. Since coding noise is an unobserved characteristic that

Table S2 Evidence in favour of BIM over PT plus additive noise

	additive noise		contextual noise	
	gains	losses	gains	losses
Zurich 03	−14084.9	−13942.6	−14085.8	−13943.9
(SE)	(46.3)	(41.7)	(46.0)	(45.9)
Zurich 06	−7824.2	−7794.7	−7820.3	−7799.6
(SE)	(33.0)	(33.1)	(33.0)	(33.1)
Beijing 05	−4768.5	−4778.0	−4777.0	−4775.7
(SE)	(31.5)	(30.8)	(31.2)	(31.1)

Test results refer to ELPD (expected log pointwise predictive density) differences, and indicate the comparative performance of the two models (Gelman et al. 2014a, Vehtari et al. 2017). Negative differences constitute evidence in favour of the BIM over PT. Tests using WAIC yield virtually identical results, and are thus not reported for parsimony.

Table S3 Model fit of BIM versus PT in L'Haridon and Vieider (2019)

	ELPD difference			ELPD difference	
country	gains	losses	country	gains	losses
Australia	−2374.5	−2147.3	Kyrgyzstan	−3783.6	−3425.8
(SE diff)	(13.3)	(12.1)	(SE diff)	(16.0)	(14.3)
Belgium	−3549.8	−3198.2	Malaysia	−2493.7	−2257.5
(SE diff)	(15.4)	(15.3)	(SE diff)	(13.5)	(12.1)
Brazil	−3277.0	−2949.3	Nicaragua	−4680.6	−4236.3
(SE diff)	(14.8)	(14.7)	(SE diff)	(17.7)	(16.2)
Cambodia	−3116.3	−2821.8	Nigeria	−7875.2	−7088.4
(SE diff)	(15.1)	(13.5)	(SE diff)	(23.1)	(20.7)
Chile	−3745.0	3361.8	Peru	−3705.2	−3346.2
(SE diff)	(15.8)	(16.7)	(SE diff)	(15.7)	(15.1)
China	−7952.8	−7191.2	Poland	−3468.2	−3128.5
(SE diff)	(23.2)	(21.6)	(SE diff)	(15.4)	(15.1)
Colombia	−4299.8	−3806.5	Russia	−2715.6	−2445.9
(SE diff)	(18.2)	(18.1)	(SE diff)	(15.3)	(14.9)
Costa Rica	−4134.9	−3725.4	Saudi Arabia	−2535.8	−2280.4
(SE diff)	(16.6)	(17.1)	(SE diff)	(18.0)	(13.3)
Czech Rep.	−3857.2	−3479.0	South Africa	−2769.4	−2499.1
(SE diff)	(16.2)	(15.7)	(SE diff)	(13.7)	(13.2)
Ethiopia	−5457.3	−4882.1	Spain	−3121.1	−2817.2
(SE diff)	(19.5)	(19.2)	(SE diff)	(14.4)	(13.9)
France	−3619.3	−3257.9	Thailand	−3078.2	−2764.3
(SE diff)	(16.5)	(16.4)	(SE diff)	(14.6)	(15.7)
Germany	−5067.6	−4574.1	Tunisia	−2872.4	−2603.4
(SE diff)	(18.4)	(18.1)	(SE diff)	(15.5)	(13.1)
Guatemala	−3264.9	−2958.4	UK	−3112.5	−2818.2
(SE diff)	(16.2)	(14.3)	(SE diff)	(15.2)	(13.7)
India	−3468.6	−3138.0	USA	−3782.7	−3385.5
(SE diff)	(15.7)	(14.4)	(SE diff)	(16.0)	(14.8)
Japan	−3274.1	−2956.4	Vietnam	−3387.6	−3064.2
(SE diff)	(14.8)	(14.4)	(SE diff)	(15.9)	(14.5)

The reported EPPD differences indicate the comparative performance of the two models (Gelman et al. 2014a, Vehtari et al. 2017). Negative differences constitute evidence in favour of the BIM over PT. All tests are significant at $p < 0.001$. Tests using WAIC yield similar results, and are thus not reported for parsimony.

drives both the discriminability parameters that are applied to choice characteristics *and* decision noise, we would expect exactly the type of omitted variable bias that is problematic for PT's causal narrative. The BIM resolved this issue by making coding noise ν observable from structural estimations.

This is exactly what we observe—in the BIM, the correlations between discriminability $\gamma \triangleq \frac{\sigma_e^2}{\sigma_e^2+\nu^2}$ and decision noise $\omega \triangleq \nu\sqrt{\alpha^2+\gamma^2}$ are even stronger. That is indeed how it should be, given that ν enters the definition of both. This also points to a way of testing this hypothesis—while we would expect discriminability $\gamma \triangleq \frac{\sigma_e^2}{\sigma_e^2+\nu^2}$ and decision noise $\omega \triangleq \nu\sqrt{\alpha^2+\gamma^2}$ to be strongly correlated, controlling for coding noise ν ought to decorrelate them, so that $p(\gamma\,|\,\nu) \perp p(\omega\,|\,\nu)$. This insight again follows textbook recommendations in causal inference, whereby the omitted variable bias is removed once the omitted variable is controlled for econometrically. One issue with this test, however, is that it again requires a rescaled version of the BIM, and that I could not make the hierarchical version of this model work on on the Zurich 2003 data. While the model converges readily on the Zurich 05 and Beijing 05 data taking about 15 minutes and delivering good convergence statistics, it takes some 2 days on the Zurich 03 data and fails to reliably estimte some parameters. This is indeed not only the case for the rescaled BIM, but even under PT I cannot recover a full covariance matrix (see also my reply below), and I need to use a brute force approach to obtain reliable estimates of the remaining parameters. I thus present a test based on 2 out of the 3 experiments.

Table S4 Parameter regressions

	ZH06 gains		ZH06 losses		BJ05 gains		BJ05 losses	
	reg 1	reg 2	reg3	reg 4	reg5	reg6	reg7	reg8
ω	-4.47^{***}	-0.86	-3.78^{***}	-0.16	-3.48^{***}	0.83	-5.25^{***}	-1.26^{*}
ν		-3.06^{***}		-3.21^{***}		-2.68^{***}		-2.58^{***}
constant	0.92	0.81	0.85	0.77	0.69	0.54	0.77	0.90

Table S4 presents a test of this hypothesis in the 2 experiments, separately for gains and losses. In each case, I regress probability-discriminability γ on the overall decision noise, ω. In a second step, I add the coding noise, ν, to the regression as a control. The results are exactly like predicted by the BIM. The correlation of γ with ω is very strong and negative as expected. Once I add ν to the regression, the regression coefficient on ω drops and becomes insignificant (except in regression 8, where it remains significant at the 10% level, but nevertheless drops substantially.) This provides a clear indication that ν indeed acts as an omitted variable, as predicted by the BIM.

Appendix S3: Separability violations under ambiguity

S3.1. Experimental details

Subjects. 48 subjects were recruited at the Melessa Lab at the University of Munich in June 2011. Only subjects who had participated in less than 3 experiments previously were invited. One subject was eliminated because she manifestly did not understand the task, and alternately chose only the sure amount or only the prospect. 38% of the subjects were male and the average age was 25 years. The experiment was run using paper and pencil.

Experimental tasks. I presented subjects with 56 different binary prospects (28 for gains, 26 for losses, and 2 mixed prospects over gains and losses). Subjects had to make a choice between these prospects and different sure amounts of money, bounded between the highest and the lowest amount in the prospect. Gains were always presented first, and losses were administered from an endowment in a second part, the instructions

for which were distributed once the first part was finished. Prospects were always kept in a fixed order. A pilot showed that this made the task less confusing for subjects, while no significant differences were found in certainty equivalents for different orders. Table S5 shows the prospects used in the usual notation $(x, p; y)$, where p indicates the probability of winning or losing x, and y obtains with a complementary probability $1 - p$.

Table S5 Decision tasks under risk and ambiguity

risky gains	uncertain gains	risky losses	uncertain losses
{0.5: 5; 0}	{0.5: 5; 0}	{0.5: –5; 0}	{0.5: –5; 0}
{0.5: 10; 0}	{0.5: 10; 0}	{0.5: –10; 0}	{0.5: –10; 0}
{0.5: 20; 0}	{0.5: 20; 0}	{0.5: –20; 0}	{0.5: –20; 0}
{0.5: 30; 0}	{0.5: 30; 0}	{0.5: –20; –5}	{0.5: –20; –5}
{0.5: 30; 10}	{0.5: 30; 10}	{0.5: –20; –10}	{0.5: –20; –10}
{0.5: 30; 20}	{0.5: 30; 20}	{0.125: – 20; 0}	{0.125: – 20; 0}
{0.125: 20; 0}	{0.125: 20; 0}	{0.125: –20; –10}	{0.125: –20; –10}
{0.125: 20; 10}	{0.125: 20; 10}	{0.25: –20; 0}	{0.25: –20; 0}
{0.25: 20; 0}	{0.25: 20; 0}	{0.385: –20; 0}	{0.385: –20; 0}
{0.385: 20; 0}	{0.385: 20; 0}	{0.625: –20; 0}	{0.625: –20; 0}
{0.625: 20; 0}	{0.625: 20; 0}	{0.75: –20; 0}	{0.75: –20; 0}
{0.75: 20; 0}	{0.75: 20; 0}	{0.875: –20; 0}	{0.875: –20; 0}
{0.875: 20; 0}	{0.875: 20; 0}	{0.875: –20; –10}	{0.875: –20; –10}
{0.875: 20; 10}	{0.875: 20; 10}	**mixed:** {0.5: 20; **–L**}	**mixed:** {0.5: 20; **–L**}

Prospects are displayed in the format $(p : x; y)$.

Notice how the exact same prospects were administered for risk (known probabilities) and uncertainty (unknown or vague probabilities). This will allow me to study ambiguity attitudes, i.e. the difference in behavior between uncertainty and risk. Preferences were elicited using choice lists, with sure amounts changing in equal steps between the extremes of the prospect.

Incentives. At the end of the game, one of the tasks was chosen for real play, and then one of the lines for which a choice had to be made in that task. This provides an incentive to reveal one's true valuation of a prospect, and is the standard way of incentivizing this sort of task. Subjects obtained a show-up fee of €4. The expected payoff for one hour of experiment was above €15.

Risk and uncertainty. Risk was implemented using an urn with 8 consecutively numbered balls. Uncertainty was also implemented using an urn with 8 balls, except that subjects were now told that, while the balls all had a number between 1 and 8, it was possible that some balls may recur repeatedly while others could be absent. The description as well as the visual display of the urns closely followed the design of Abdellaoui et al. (2011). The main differences were that I ran the experiment using paper and pencil instead of with computers; that I used numbers instead of colours in order to allow for black and white printing; and that I ran the experiment in sessions of 15-25 subjects instead of individually.

The decision model closely follows the one for risk used above, but with all parameters doubled for ambiguity. Ambiguity parameters are coded as deviations of the equivalent parameters for risk, mostly to

increase computational efficiency and without loss of generality. All estimations are executed directly using the density around the switching point. The model used for PT looks as follows (@add code upon publication):

Figure S14 furthermore shows the distributions of the original BIM parameters (the same caveat as above on the rescaling of the noise applies here, too). Panel A shows coding noise for risk and ambiguity, where coding noise can be clearly seen to be larger under ambiguity. Panel B shows the two standard deviations for the priors. The orders of magnitude shown are indeed consistent with the examples I gave above. Panels C and D show the resulting likelihood-discriminability and outcome-discriminability parameter of the BIM, respectively. The similarity of panel C with figure 2a in the main paper bears again testimony for the (almost) straight mapping from one into the other.

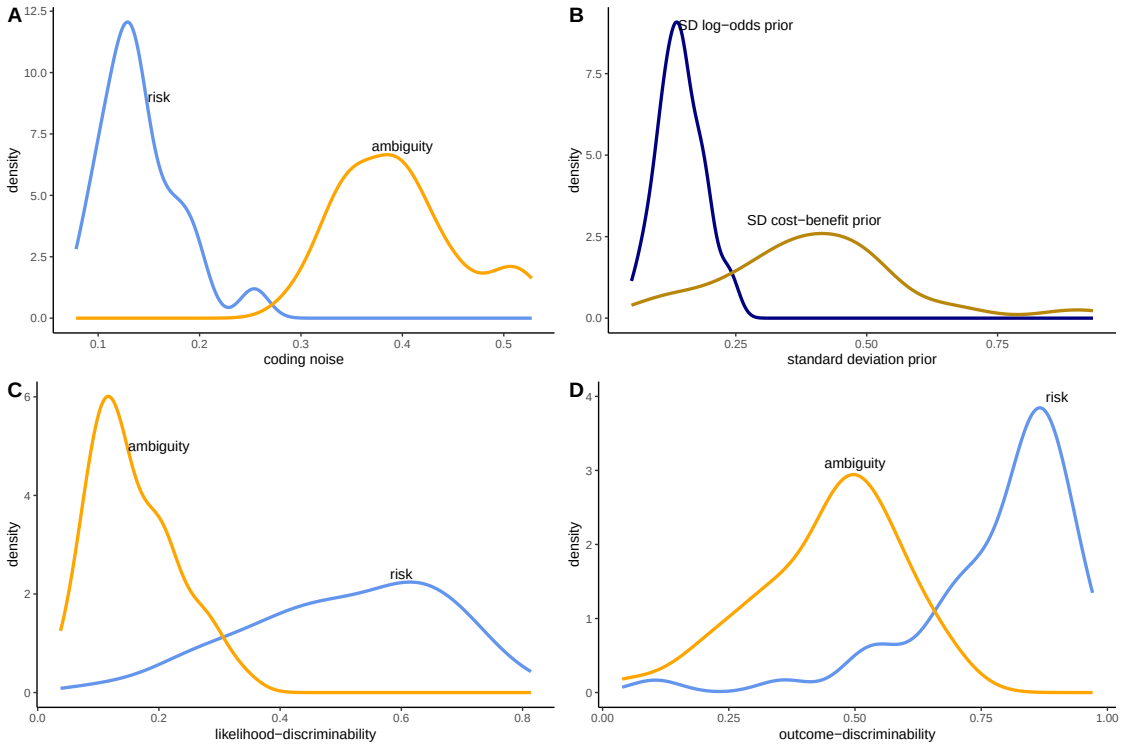


Figure S14 BIM parameters for ambiguity data

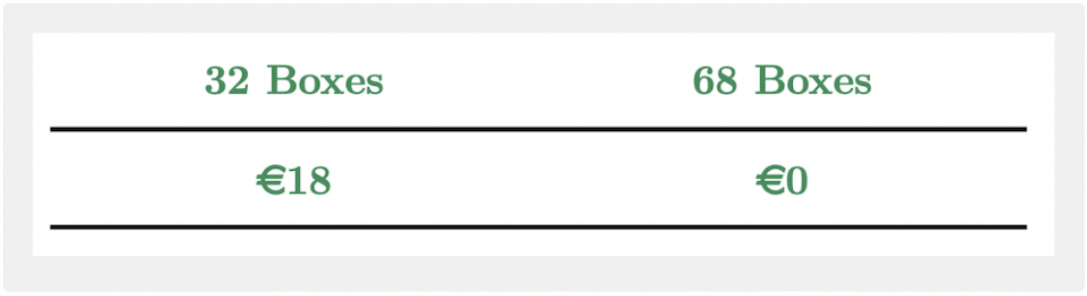
Appendix S4: Instructions risky choice versus mirrors

First, all subjects saw the following general instructions:

Welcome to this experiment in decision making. Below, you will be asked to repeatedly choose between different options. In the end, **we will select 10 students to play one randomly chosen decision for real money**. Please read these instructions carefully, and pay close attention to the options being presented to you. Your final payoffs may depend on this.

In each task, you will be presented with 2 options. In both options, your payout will depend on **100 BOXES containing different monetary amounts**. The two sets of boxes will be displayed like in the example below. The counts of boxes are listed at the top, and below them you can see the monetary amounts contained in those boxes. In this example, the option at the top has 32 Boxes containing €18, and 68 Boxes containing €0 (nothing). The option at the bottom has 100 Boxes containing €6.

Please make a choice



Notice that the order of the boxes, the numbers of boxes, and associated amounts will change from task to task. **Please pay close attention to these dimensions before indicating your choice**. Please play careful attention to these instructions. After reading the instructions, you will be asked some comprehension questions. **You will only be allowed to proceed to the experiment if you correctly answer all comprehension questions**.

Subjects assigned to the RANDOM BOX treatment then saw the following instructions:

RANDOM BOX

I the upcoming tasks, **we will RANDOMLY DRAW one of the 100 boxes** from whichever option you have chosen to determine your payoff. All the boxes are equally likely to be drawn.

Example: Take a look at the example below. If you have selected the option at the top, we will make a **RANDOM DRAW** from a bag with numbers from 1 to 100. If the number drawn falls in the range 1–58, you get €32. If it falls in the range 59-100, you get € 0 (nothing). If you have selected the option at the bottom, you will simply be paid the sure amount of € 12, since all 100 boxes contain that same amount.

Please make a choice:

58 Boxes	42 Boxes
€32	€0

100 Boxes
€12

Subjects assigned to the AVERAGE BOX treatment instead saw the following instructions:

AVERAGE BOX

In the upcoming tasks, **we will pay you by calculating the AVERAGE amount of money across all 100 Boxes** from whichever option you have chosen. That is, we will add up the amount from each of the 100 boxes and divide the sum by 100.

Example: Take a look at the example below. If you have selected the option at the top, we will add (58 * €32 + 42 * €0)/100 = € 18.56. If you have selected the option at the bottom, you will simply be paid €12, since (€12 * 100)/100 = €12.

Please make a choice:

58 Boxes	42 Boxes
€32	€0

100 Boxes
€12

All subjects had to answer 6 comprehension questions as the ones shown below. The questions served the purpose of further emphasizing the treatment:

Suppose that the choice above determines your payoff, and that **you have chosen the option at the top**. What is the chance that you will obtain €18?

- 0 in 100 (0%)
- 50 in 100 (50%)
- 100 in 100 (100%)

Suppose that the choice above determines your payoff, and that **you have chosen the option at the top**. What is the chance that you will obtain €9?

- 0 in 100 (0%)
- 50 in 100 (50%)
- 100 in 100 (100%)