Valuations Under Tradeoff Complexity*

Cassidy Shubatt Jeffrey Yang

First posted: May 21, 2026
Updated: May 21, 2026

Abstract

Valuation tasks are a workhorse method for testing theories of individual preferences. However, a body of evidence suggests that complexity produces systematic measurement error in valuations, which raises the question of how researchers should interpret and utilize valuation data. To formally study this issue, we develop a model of complexity-driven noise in the valuation of risky prospects, in which the difficulty of comparing options to prices produces systematic noise in their valuations. We show how this model of noise can explain a number of documented valuation patterns that are difficult to rationalize under prevailing theories of risk preferences, as well as novel experimental evidence of systematic inconsistencies across valuation formats. We then characterize which valuation-based tests of preferences are robust to complexity in our model. While complexity distorts the levels of valuations in our model, differences between valuations can be informative, so long as complexity is held fixed across valuation tasks. We provide a formal criterion for robustness and apply it to valuation designs in the literature.

Keywords: Complexity, choice under risk, experiments

*Shubatt: Department of Economics, Harvard University. cshubatt@g.harvard.edu. Yang: Department of Economics, University of California, Santa Barbara. jeffrey.yang97@gmail.com. We thank Benjamin Enke, Erik Eyster, David Laibson, Daniel Martin, Joshua Schwartzstein, and Jason Somerville for helpful comments and suggestions.

1 Introduction

Valuation tasks are a workhorse method for measuring preferences and testing theories of individual decision-making. Valuations have played a central role in testing and developing theories of choice under risk (Tversky and Kahneman, 1992; Holt and Laury, 2002; Andersen et al., 2008; Andreoni and Sprenger, 2011; Bernheim and Sprenger, 2020), among other choice domains, and they are heavily featured in recent experimental methodologies used to study preferences (McGranaghan et al., 2024a,b). Outside the laboratory, valuation data are commonly collected in applied work, where they are used to measure willingness to pay, estimate preference parameters, and evaluate interventions in field and development settings (see Jack et al., 2022, for a review).

At the same time, a body of evidence suggests that elicited valuations are subject to systematic measurement error that complicates their interpretation (Harrison et al., 2005; Andersen et al., 2006; Beauchamp et al., 2020; Freeman and Mayraz, 2019; Freeman et al., 2019; Jack et al., 2022; Holden et al., 2025). Implicit in many of these accounts is the idea that these distortions arise due to subjects' uncertainty over or difficulty with forming valuations, consistent with models of how cognitive frictions produce forms of systematic noise in valuations (Enke and Graeber, 2023; Frydman and Jin, 2025; Shubatt and Yang, 2024). Notably, recent work has argued that the canonical fourfold pattern in the valuation of risky prospects is driven by complexity-driven noise as opposed to underlying preferences (Enke and Graeber, 2023; Oprea, 2024), a possible explanation for why the fourfold pattern is less pronounced in direct choice between prospects than in valuations (Bouchouicha et al., 2024; Imai et al., 2025).

This raises several questions. Which empirical patterns in valuations should be interpreted as reflecting preferences as opposed to artifacts of complexity? Which experimental tests involving valuation data are valid in the face of complexity-driven noise? More generally, how should valuation-based tests be designed if the goal is to learn about underlying risk preferences? Without precise formal structure on how complexity can distort valuations, it is difficult to provide sharp answers to these questions. In this paper, we develop a model of systematic complexity-driven noise in the valuation of risky prospects, and use it to reconcile documented puzzles in the extant literature, characterize what valuation data can tell us about preferences, and derive implications on which valuation tests are robust to complexity-driven noise.

Motivating evidence for systematic noise. To motivate this exercise, we first conduct an experiment designed to demonstrate the importance of systematic noise in valuations. In our experiment, subjects report valuations for the same set of lotteries using four distinct multiple price list variants over which the numeraire good varies. Under any account of stable preferences and unbiased noise in valuations, the preference ordering implied by subjects' valuations of these lotteries should be invariant across these alternative valuation modes. We strongly reject this implication. Each of the four valuation modes produces a distinct modal preference ordering over the same set of lotteries. Our data are difficult to reconcile with the assumption of mean-zero noise in valuations, even under flexible assumptions on heterogeneity in preferences and decision noise. These results strongly suggest that noise in valuations is systematically biased, and motivate models that put quantitative structure on this bias.

Modeling valuations under tradeoff complexity. We model a decision-maker who values a lottery x against a price, or numeraire good—formally, an ordered continuum of numeraire lotteries $\mathcal{Z}$—on the basis of potentially imprecise comparisons between x and prices in $\mathcal{Z}$. Formally, the DM learns about the valuation of x through a noisy signal of how x ranks against prices in $\mathcal{Z}$, which are less precise for prices that are harder to compare to x. This process results in a posterior range of uncertainty over x's valuation in terms of $\mathcal{Z}$, which is more likely to include prices that are hard to compare to x.

To put structure on the ease of comparison, we adapt the measure of tradeoff complexity developed in Shubatt and Yang (2024), in which the difficulty of comparing lotteries increases as the DM faces greater tradeoffs across their distributions. Our resulting model predicts systematic noise that compresses valuations toward the range of prices that are hard to compare to the lottery. This straightforwardly explains the pattern of valuation inconsistencies we document in our experiment, as well as related inconsistencies, such as the reversal of the fourfold pattern under alternative valuation methods (Feldman and Ferraro, 2024; Shubatt and Yang, 2024).[1]

Reconciling valuation patterns. We then confront our model with a set of valuation patterns documented in the extant literature that are difficult to rationalize under prevailing theories of risk preferences. Our model can rationalize 1) inverse-S weighting in certainty equivalents and the absence of an aggregate common ratio effect in compounded certainty equivalents in McGranaghan et al. (2024b); 2) the presence of non-linear probability weighting and the absence of rank-dependence in the tests of Bernheim and Sprenger (2020); 3) inverse-S weighting in certainty equivalents and the absence of independence violations away from certainty in uncertainty equivalents in Andreoni and Sprenger (2011); and 4) boundary effects in certainty equivalents documented in Kontek (2018). Whereas leading theories of risk preferences, such

[1] Shubatt and Yang (2024) show that this pattern can be rationalized in a discrete model of valuations in which the DM chooses in a multiple price list based on noisy signals of value comparisons. Our model can be derived as a limit case of this model under an alternative signal structure.

as expected utility theory and rank-dependent utility, would require substantive modifications to accommodate these valuation patterns under standard assumptions on noise, we show that under expected utility, our specification of noise can rationalize key features of these patterns.

Implications for valuation-based tests of preferences. In the exercise above, there are features of the data that cannot be rationalized by our model of expected utility preferences and complexity-driven noise, which highlights areas in which our assumed theory of expected utility may be inadequate. Motivated by this, we use our model to derive implications for what can be learned about preferences from valuation data in the presence of complexity-driven noise, and for which valuation-based tests of preferences are robust to this form of noise.

Our model implies that while the *levels* of valuations are in general distorted by systematic noise, and should therefore be interpreted with caution, *differences* between valuations can be informative about underlying preferences, so long as tradeoff difficulty is held constant across valuation tasks—since this ensures complexity-driven noise is held fixed across tasks, differences in valuations are informative of preferences. We show that many common uses of valuation tasks do not satisfy this criterion and are thus subject to complexity-driven confounds, such as the use of certainty equivalents to estimate risk aversion or measure the fourfold pattern. At the same time, a number of recently proposed valuation methodologies satisfy this criterion, such as the paired valuation tasks in McGranaghan et al. (2024b,a) used to test for common ratio, common consequence, and mixture preferences, the test for rank-dependence in Bernheim and Sprenger (2020), and the uncertainty equivalents method in Andreoni and Sprenger (2011) used to test for independence violations.

Contributions and related literature. Our paper relates to a longstanding literature that uses valuation data to study risk preferences (e.g. Tversky and Kahneman, 1992; Holt and Laury, 2002; Andersen et al., 2008; Andreoni and Sprenger, 2011; Bernheim and Sprenger, 2020; McGranaghan et al., 2024a,b). The predominant approach in this literature has been to interpret elicited valuations under the assumption of unbiased noise, and to use the resulting valuation data to inform and modify theories of preferences. Our paper takes a complementary approach: we revisit standard assumptions on noise in valuations, and focus on modeling systematic, complexity-driven noise. In doing so, we show that a number of puzzling valuation patterns in the literature can be rationalized by our model of noise, assuming standard risk preferences, and derive implications for which valuation-based tests of preferences are robust to complexity-driven noise.

In characterizing complexity-robust valuation designs, our paper relates to a broader methodological discussion on measurement error in valuation elicitations. A body of work has shown that elicited valuations are subject to various forms of systematic measurement error, such as pull-to-center biases and range effects (Harrison et al., 2005; Andersen et al., 2006; Beauchamp et al., 2020; Freeman and Mayraz, 2019; Freeman et al., 2019; Jack et al., 2022). To deal with these concerns, past work can be broadly categorized into two approaches. The first approach corrects for systematic errors at the "analysis stage", assuming a parametric structure on these distortions and uses econometric corrections to recover latent preferences from valuations (e.g. Andersen et al., 2006; Beauchamp et al., 2020; Jack et al., 2022). The second approach seeks to control for these errors at the "implementation stage", comparing and developing methods of eliciting a given valuation so as to reduce the influence of subject confusion (e.g. Andersen et al., 2006, 2008; Bruner, 2011; Hardisty et al., 2013; Cason and Plott, 2014; Holden et al., 2025), as well as studying the theoretical properties of elicitation methods (e.g. Karni and Safra, 1987; Azrieli et al., 2018).

Our paper takes a third approach: we seek to address systematic noise at the "design stage", using our model to provide guidance on *which* valuations to elicit in the face of complexity-driven noise. We view this approach as complementary. Relative to the first approach, our results on robust valuation designs do not rely on specific parametric assumptions but instead only the general structure of our noise specification, at the cost of exact preference recovery. Relative to the second approach, we abstract from differences in framing and subject confusion across different methods of implementing a valuation task, and instead focus on modeling an underlying source of difficulty in forming valuations—tradeoffs in comparing options against prices—that likely applies across different implementations.[2]

Finally, our paper relates to a broader literature studying the implications of noise for the design and interpretation of revealed preference measures in experiments (Harless and Camerer, 1994; Ballinger and Wilcox, 1997; Loomes and Sugden, 1998; McGranaghan et al., 2024b; Echenique and Tserenjigmid, 2026), which has considered a broader set of revealed preference measures beyond valuations, such as direct choice.[3] Our focus on modeling noise in valuations is motivated by their wide use in experimental research, as well as their appealing properties for testing preference theories (see e.g. Bernheim and Sprenger, 2020; McGranaghan et al., 2024a; O'Donoghue and Somerville, 2026).

This paper proceeds as follows. Section 2 presents motivating evidence for systematic noise

[2]For instance, (Enke and Graeber, 2023) document that subjects' subjective uncertainty over their valuations of lotteries predict a stronger pull-to-center effect in both stated valuations and multiple price lists. Likewise, preference reversals between valuation and direct choice in lottery choice, which Shubatt and Yang (2024) show can be attributed to tradeoff complexity, have been documented in valuations elicited both using the Becker-deGroot-Marschak mechanism and multiple price lists.

[3]See McGranaghan et al. (2024a) and Echenique and Tserenjigmid (2026) for a recent discussion of the implications of noise for valuation- and choice-based preference tests

in valuations. Section 3 develops our model of complexity-driven noise in valuations. Section 4 applies the model to explain documented patterns in the valuation literature, and Section 5 develops implications for the design of complexity-robust valuation-based tests. Section 6 concludes. Proofs are relegated to the Appendix.

2 Evidence for Systematic Noise in Valuations

When interpreting valuation data, a standard assumption is that noise in valuations is unbiased. We conduct an experiment to demonstrate the presence and quantitative importance of systematic noise in valuations.

To fix ideas, consider the following measurement framework. Denote by $v_{\mathcal{Z}}(x)$ a decision-maker's true valuation of a lottery x against the price structure $\mathcal{Z}$. Here, $\mathcal{Z}$ describes the numeraire against which x is valued, i.e. whether x is valued in terms of certain payments or probability equivalents. We assume that true valuations are consistent across price structures, i.e. $v_{\mathcal{Z}}(x) > v_{\mathcal{Z}}(y)$ implies $v_{\mathcal{Z}'}(x) > v_{\mathcal{Z}'}(y)$.[4]

Let $V_{\mathcal{Z}}(x)$ denote the experimenter's measurement of $v_{\mathcal{Z}}(x)$, and let $\epsilon_{\mathcal{Z}}(x) \equiv V_{\mathcal{Z}}(x) - v_{\mathcal{Z}}(x)$ denote the valuation error. A standard assumption is that these errors are unbiased, so that comparisons of elicited valuations are informative of an individual's true valuations. We consider two forms of this *unbiased noise* assumption:

A1. For all $x, y, \mathcal{Z}$, $E[V_{\mathcal{Z}}(x)] > E[V_{\mathcal{Z}}(y)] \iff v_{\mathcal{Z}}(x) > v_{\mathcal{Z}}(y)$.

A2. For all $x, y, \mathcal{Z}$, $\mathbb{P}(V_{\mathcal{Z}}(x) > V_{\mathcal{Z}}(y)) > 0.5 \iff v_{\mathcal{Z}}(x) > v_{\mathcal{Z}}(y)$.

These assumptions are implied by familiar error specifications. For instance, mean-zero errors imply A1, median-zero differences $\epsilon_{\mathcal{Z}}(x) - \epsilon_{\mathcal{Z}}(y)$ imply A2, and mean-zero and symmetric errors imply both.

An immediate implication of unbiased noise is that a decision-maker's valuations must be consistent across price structures. A1 and A2 imply, respectively,

Implication 1. $\mathbb{E}[V_{\mathcal{Z}}(x)] > \mathbb{E}[V_{\mathcal{Z}}(y)] \iff \mathbb{E}[V_{\mathcal{Z}'}(x)] > \mathbb{E}[V_{\mathcal{Z}'}(y)]$

Implication 2. $\mathbb{P}(V_{\mathcal{Z}}(x) > V_{\mathcal{Z}}(y)) > 0.5 \iff \mathbb{P}(V_{\mathcal{Z}'}(x) > V_{\mathcal{Z}'}(y)) > 0.5$

We now discuss an experiment that tests for this consistency.

[4] In our formal treatment of valuation tasks in Section 3, we show that this form of consistency across price formats must hold if underlying preferences are monotonic.

2.1 Design

Subjects value four pairs of risky prospects that pay an amount of money with some probability. Each pair consists of lotteries with the same expected value; one lottery (the Risky lottery, denoted ℓ_R) is a mean-preserving spread of the other (the Safe lottery, denoted ℓ_S). Each pair is identified by a value $p \in \{0.1, 0.2, 0.8, 0.9\}$: the Risky lottery pays \$30 with $20 \cdot p$% chance, and the Safe lottery pays $\$30 \cdot p$ with 20% chance.

For each pair, we elicit subjects' valuations of the lotteries using four valuation modes:

1. Money equivalents a) (M-a): value options against \$$m$ with 20% chance
2. Money equivalents b) (M-b): value options against \$$m$ with 50% chance
3. Probability equivalents a) (P-a): value options against r% chance of \$30
4. Probability equivalents b) (P-b): value options against r% chance of \$40

Notice that if the subjects have monotonic preferences over lotteries, their true valuations over lottery pairs should be ordinally consistent across these valuation modes. We are interested in assessing whether observed valuations satisfy this consistency.

Since the M-a valuations of the Safe lotteries and likewise the P-a valuations of the Risky lotteries are trivial, we infer subjects' valuations in those cases, rather than elicit them directly. That is, we assume that subjects' M-a valuation of \$$y$ with 20% chance is equal to y, and likewise that subjects' P-a valuation of \$30 with q% chance is equal to q.

In our experiment, we elicit these valuations using a multiple price list (MPL) elicitation. Subjects completed money-equivalent and probability-equivalent valuation tasks in separate blocks, with both the order of price lists within each block and the order of blocks randomized. Each subject had a 1/5 chance for their decisions to be incentivized; if this occurred, we randomly selected a single price list in the experiment, and their bonus was determined by the lottery they chose in a randomly selected row of that price list. See Appendix C for more design details and screenshots of the experimental interface.

2.2 Results

252 subjects were recruited on Prolific to complete the study. The study took 23.2 minutes on average. Subjects earned a show-up payment of \$3.50, and an average bonus of \$0.87.

Figure 1, which plots the mean lottery valuations in each pair under each of our four modes of valuation, reveals clear aggregate differences across the modes of valuation. Panel a) shows that in M-a valuations, subjects on average value ℓ_R above ℓ_S for $p \in \{0.1, 0.2\}$, and value ℓ_R

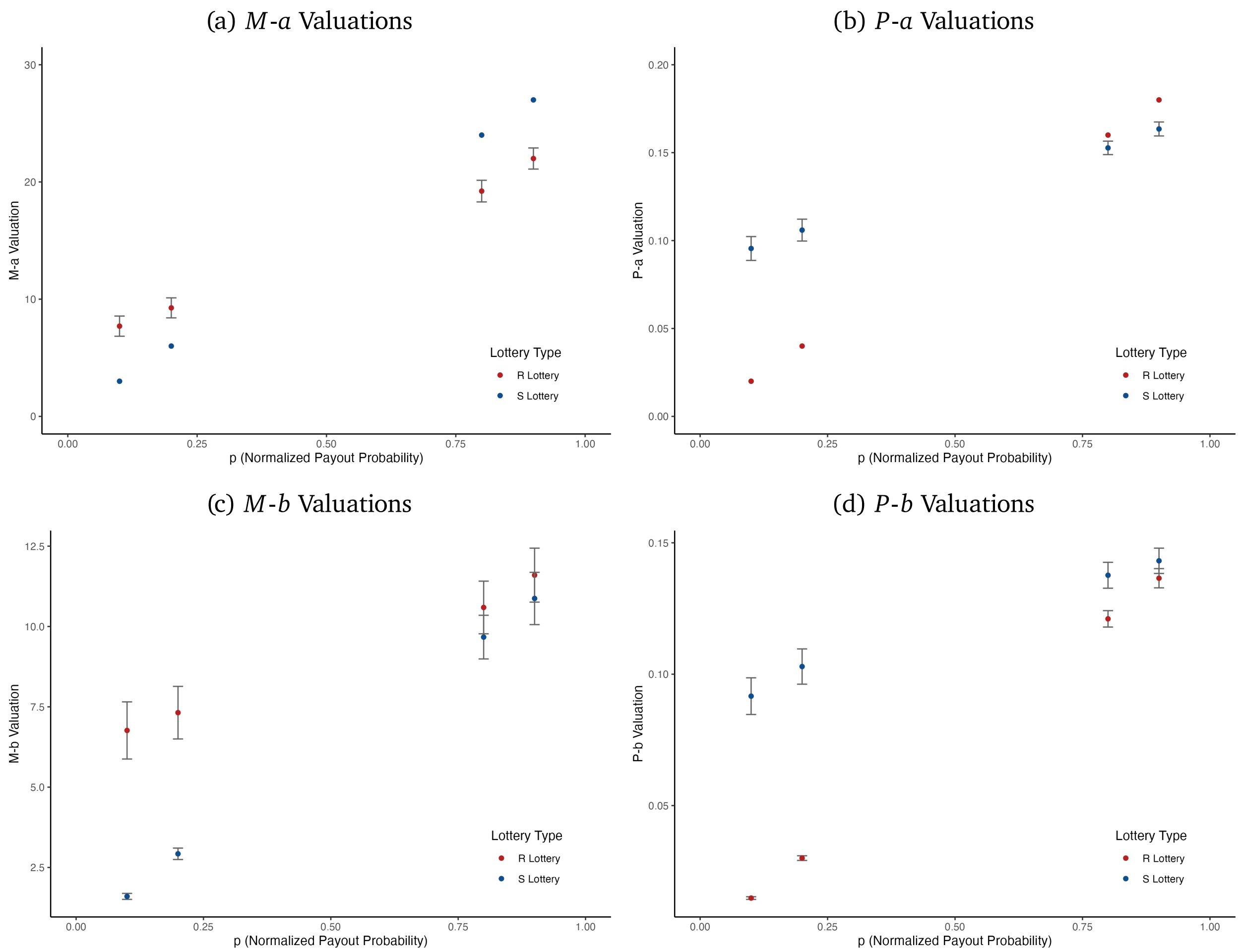

Figure 1: Valuation levels. Whiskers plot 95% confidence intervals.

below ℓ_S for $p \in \{0.8, 0.9\}$. That is, M-a valuations are on average risk-seeking for low p, and risk-seeking for high p. This is consistent with McGranaghan et al. (2024), who document the same pattern of aggregate valuations in similar MPL valuation tasks.

Turning to panel b), however, we see the *opposite* pattern in P-a valuations. Subjects on average value the ℓ_R *below* ℓ_S for $p \in \{0.1, 0.2\}$, and value ℓ_R *above* ℓ_S for $p \in \{0.8, 0.9\}$. That is, P-a valuations are on average risk-averse for low p, and risk-seeking for high p. Panels c) and d) show M-b and P-b additionally exhibit inconsistent aggregate patterns: M-b valuations are weakly risk-seeking for all p, and P-b valuations are weakly risk-averse for all p. In sum, each of our four valuation modes leads to different conclusions about the average ranking between the paired lotteries.

We can more directly study apparent inconsistencies in risk preferences across our four valuation modes by examining, for each lottery pair, the proportion of subjects whose valuations indicate a preference for the Risky vs. Safe lottery (that is, the proportion of risk-seeking valu-

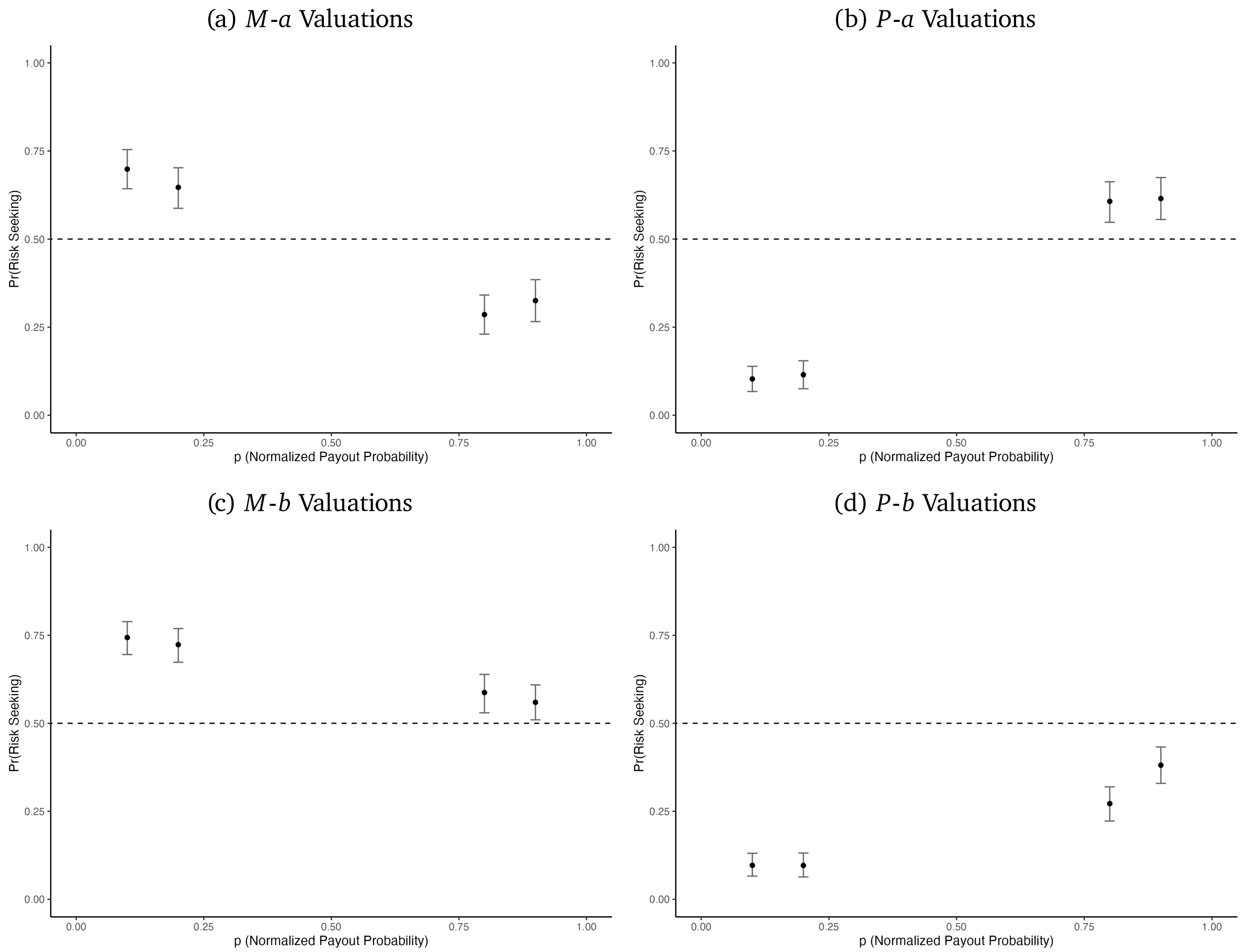

Figure 2: Proportion of risk-seeking valuations. Whiskers plot 95% confidence intervals.

ations) under each valuation mode.

Figure 2 plots these proportions. The patterns that emerge are largely consistent with Figure 1. When lotteries are valued under M-a, the majority prefers ℓ_R (risk-seeking) for low p and prefers ℓ_S (risk-averse) for high p. When the same lotteries are valued under P-a, we observe the opposite modal preference: the majority prefers ℓ_S (risk-averse) for low p and prefers ℓ_R (risk-seeking) for high p. Similarly, under M-b valuations, the majority prefers ℓ_R (risk-seeking) for all p, and under P-b valuations, the majority prefers ℓ_S (risk-averse) for all p. Table 1 summarizes the aggregate valuation patterns.

These aggregate results are inconsistent with Implications 1 and 2 of unbiased noise, assuming the data are generated by a representative agent. However, these results could be explained under unbiased noise if we allow for heterogeneity in preferences and valuation noise across individuals, under certain forms of correlations between preferences and noise.[5] As such, we

[5]See (McGranaghan et al., 2024b) for a discussion of how heterogeneity can lead to apparent inconsistencies

Table 1: Summary of aggregate valuation patterns.

	$p \in \{0.1, 0.2\}$	$p \in \{0.8, 0.9\}$
M-a valuations:	$\ell_R \succ \ell_S$ (risk-seeking)	$\ell_R \prec \ell_S$ (risk-averse)
P-a valuations:	$\ell_R \prec \ell_S$ (risk-averse)	$\ell_R \succ \ell_S$ (risk-seeking)
M-b valuations:	$\ell_R \succ \ell_S$ (risk-seeking)	$\ell_R \succ \ell_S$ (risk-seeking)
P-b valuations:	$\ell_R \prec \ell_S$ (risk-averse)	$\ell_R \prec \ell_S$ (risk-averse)

turn to within-subjects evidence of inconsistencies across valuation modes. Column (1) of Table 2 reports, for each pair of valuation modes and lottery pair, the proportion of subjects who exhibit inconsistent valuations across the two modes. In every comparison for which we observe an aggregate reversal, the majority of subjects make inconsistent valuations: each of these inconsistency rates are weakly greater than 50%, and we can reject the null of equality in half of the relevant comparisons. This is inconsistent with Implication 2 of unbiased noise under *any* specification of heterogeneity in preferences and valuation noise across individuals, so long as noise is independent across valuations.[6]

These within-subject reversals are not only prevalent in our sample, but are also systematic in nature. Column (2) of Table 2 reports the proportion of subjects who exhibit an inconsistency in the direction that *coheres* with the aggregate patterns summarized in Table 1. For instance, in the comparison of M-a and P-a valuations, an inconsistency that coheres with the aggregate pattern would assign a higher valuation to ℓ_R under M-a and a higher valuation to ℓ_S under P-a for $p \in \{0.1, 0.2\}$, and vice versa for $p \in \{0.8, 0.9\}$. A comparison of columns (1) and (2) reveals that for each relevant comparison, the majority of inconsistencies directionally cohere with the aggregate patterns.

2.3 Interpretation and Discussion

In our experiment, valuations exhibit markedly different patterns across multiple ordinally equivalent valuation formats. Under the maintained assumption that underlying preferences are consistent across valuation formats, these results strongly reject the assumption of unbiased noise in valuations.

To account for these results, one must either relax the assumption that preferences are consistent across valuation formats, or relax the assumption of unbiased noise. While past work has taken the former approach, showing that models of reference-dependent preferences (Sprenger, 2015; Feldman and Ferraro, 2024) and salience (Bordalo et al., 2012) can explain certain in-

in aggregated preferences under unbiased noise.

[6]To see this, note that under Implication 2, the probability that an individual makes an inconsistent valuation takes the form $(1-p_A)p_B + (1-p_B)p_A$, where $p_A, p_B \geq 0.5$; this probability is weakly less than 0.5.

Table 2: Individual-Level Valuation Inconsistencies

	Pr(Incons.)	Pr(Incons.) Predicted	Pr(Incons.)> 0.5 (t-statistic)
	(1)	(2)	(3)
Panel A. M-a vs. P-a			
$p = 0.1$	0.66	0.63	5.30***
$p = 0.2$	0.59	0.56	2.81**
$p = 0.8$	0.57	0.44	2.16*
$p = 0.9$	0.58	0.44	2.68**
Panel B. M-b vs. P-b			
$p = 0.1$	0.73	0.69	9.67***
$p = 0.2$	0.70	0.66	7.94***
$p = 0.8$	0.54	0.43	1.41
$p = 0.9$	0.53	0.35	1.18
Panel c. M-a vs. M-b			
$p = 0.8$	0.53	0.41	1.03
$p = 0.9$	0.53	0.38	1.22
Panel d. P-a vs. P-b			
$p = 0.8$	0.54	0.44	1.32
$p = 0.9$	0.54	0.38	1.28

Column (1) reports the proportion of valuations for each ℓ_R, ℓ_S pair that are inconsistent across valuation formats. Column (2) reports the proportion of valuations for each ℓ_R, ℓ_S pair that are inconsistent and that coheres with the aggregate patterns summarized in Table 1. Column (3) reports t-statistics for a two-sided bootstrapped t-test. Stars indicate significance level of the test. $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

consistencies across valuation formats, these models cannot account for the patterns of inconsistencies that we document, as we discuss in Appendix A.6.

We take the complementary approach of assuming stable preferences, and instead focus on the possibility of systematic noise in valuations. In tandem with a body of existing work arguing that complexity produces systematic errors in valuations, this raises a basic question: which valuation patterns are attributable to underlying preferences, as opposed to systematic noise? Addressing this question requires a model of stochastic valuations that is more descriptive than unbiased noise. We develop one such model below.

3 Modeling Systematic Noise in Valuations

3.1 Valuation Framework

Let X denote the set of finite monetary lotteries, where each $x \in X$ is described by a probability mass function $f_x : \mathbb{R} \to [0,1]$ where $f_x(w) > 0$ for finitely many w; let F_x^{-1} denote the associated quantile function. Let $>_{FOSD}$ denote the first-order stochastic dominance relation on X. Consider a decision-maker (DM) tasked with valuing lotteries. The DM's preferences are described by a continuous weak order $\succeq$ on X that is monotonic, i.e. $x >_{FOSD} y \implies x \succ y$; say that $\succeq$ is *proper* if it satisfies these properties.

We formalize valuation tasks as follows. For $B \subseteq \mathbb{R}$, let $\mathcal{I}(B)$ denote the set of all non-degenerate sub-intervals of B. For $\mathcal{W} \in \mathcal{I}(\mathbb{R})$, a *price structure* is a family of lotteries $\mathcal{Z} = \{z_w\}_{w \in \mathcal{W}}$ where $z_w <_{FOSD} z_{w'}$ if and only if $w < w'$, and the mapping $w \mapsto z_w$ is continuous. The index set $\mathcal{W}$ has the interpretation of the numeraire value, i.e. the monetary amount in a certainty equivalents task or the payout probability in a probability equivalents task.

A *valuation task* $(x, \mathcal{Z})$ is a pairing of a lottery x and price structure $\mathcal{Z}$ such that $z \preceq x \preceq z'$ for some $z, z' \in \mathcal{Z}$. By continuity and monotonicity of $\succeq$, this implies that there exists a unique $z \in \mathcal{Z}$ that is indifferent to x; denote by $v_{\mathcal{Z}}(x)$ the index of this lottery, which has the interpretation of the DM's true valuation of x in terms of $\mathcal{Z}$. Denote by $\underline{w}_{\mathcal{Z}}(x) \equiv \inf\{w \in \mathcal{W} : z_w \not<_{FOSD} x\}$ and $\overline{w}_{\mathcal{Z}}(x) \equiv \sup\{w \in \mathcal{W} : z_w \not>_{FOSD} x\}$ the endpoints of the range of undominated prices in $\mathcal{W}$, and call a valuation task $(x, \mathcal{Z})$ *bounded* if $z \leq_{FOSD} x \leq_{FOSD} z'$ for some $z, z' \in \mathcal{Z}$.

Example 1. (Examples of Valuation Tasks). Here, let $(a, b, .., c; p, q, ..., r)$ denote the lottery that pays out $a, b, ..., c \neq 0$ with probabilities $p, q, ..., r$, respectively, and 0 otherwise.

- *Certainty equivalents.* Here, the DM values x against certain payments. This is described by $(x, \mathcal{Z})$, where $\mathcal{Z} = \{(w; 1)\}_{w \in \mathbb{R}}$.

- *Probability equivalents.* Here, the DM assesses the payoff probability p that makes $(m; p)$ indifferent to x, where m is weakly greater than the upper bound of the support of x. This is described by $(x, \mathcal{Z})$, where $\mathcal{Z} = \{(m; w)\}_{w \in [0,1]}$.

- *Money equivalents.* Here, the DM states the monetary m amount that would make $(m; p)$ indifferent to x. This is described by $(x, \mathcal{Z})$, where $\mathcal{Z} = \{(w; p)\}_{w \in \mathbb{R}}$.

- *Equalizing reductions.* Given a lottery $x = (a, b, c; p, q, 1-p-q)$, the DM states the reduction in payoff c that would be needed to compensate a given increase k in payoff b. This is described by $(x, \mathcal{Z})$, where $\mathcal{Z} = \{(a, b+k, c-w; p, q, 1-p-q)\}_{w \in \mathbb{R}}$.

3.2 Modeling Complexity Driven Noise

Rather than assuming the DM has perfect access to her preferences and therefore her valuations, we model a DM who values options on the basis of imprecise comparisons between x and the numeraire, motivated by the idea that decision-makers may find some options more difficult to compare than others (Natenzon, 2019; Shubatt and Yang, 2024). For instance, in assessing the certainty equivalent of $x = (\$10; 0.3)$, the DM presumably understands that x should be valued less than \$10 but may have an imprecise understanding of how x should be valued above or below \$2. We formalize how such imprecision produces systematic noise in valuation in the case of bounded valuation tasks, and later extend the model to general valuation tasks.

Signal structure. Consider a DM who faces a bounded valuation task $(x, \mathcal{Z})$. Rather than directly observing her true valuation $v_{\mathcal{Z}}(x)$, the DM learns about it through a set of noisy local comparisons between x and the prices in $\mathcal{Z}$. Formally, we model valuations as the outcome of Bayesian learning. The DM begins with uncertainty over $v_{\mathcal{Z}}(x)$, with priors uniformly distributed on $\mathcal{W}$. She receives a signal $s \in S \equiv \{-1, 0, +1\}^{\mathcal{W}}$ that summarizes a collection of local comparisons between x and prices in $\mathcal{Z}$, forms a posterior belief over $v_{\mathcal{Z}}(x)$, and then reports her posterior expectation $V_{\mathcal{Z}}(x)$. The likelihood function $\pi(s|v)$, which describes the distribution of signals conditional on the true valuation $v \in \mathcal{W}$ is characterized by two properties.

First, for each price $w \in \mathcal{W}$, the local signal $s(w) \in \{-1, 0, +1\}$ either reveals that x is worth less than z_w ($s(w) = -1$), that x is worth more than z_w ($s(w) = +1$), or is uninformative about the comparison ($s(w) = 0$):

1. $\pi(\{s \in S : \exists w \leq v \text{ s.t. } s(w) = 1\}|v) = \pi(\{s \in S : \exists w \geq v \text{ s.t. } s(w) = -1\}|v) = 0.$

Second, the arrival of information about the comparison between x and each z_w is independent across w, and informative local signals arrive more frequently in regions of prices that are easier to compare to x:

2. Let $N(I; s) = \sum_{w \in I} |s(w)|$ count the number of informative local signals produced by s in an interval $I \in \mathcal{I}(\mathcal{W})$. Under $\pi(s|v)$, $N(I; s)$ is a Poisson process with the rate function

$$\lambda^x_{\mathcal{Z}}(t) = -\frac{1}{\overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)} \cdot \ln(1 - \tau(x, z_t)),$$

where $\tau : X \times X \to [0, 1]$ is a *comparability function*, which satisfies, for all $x, y \in X$, $\tau(x, y) = \tau(y, x)$, $\tau(x, y) = 0$ whenever $x \sim y$, and $\tau(x, y) = 1$ whenever $x >_{FOSD} y$.

Here, $\tau(\cdot,\cdot)$ describes the ease of comparing lotteries: if $\tau(x,z_w) = 1$ for a neighborhood of prices, the DM perfectly learns how x compares against that neighborhood; if instead $\tau(x,z_w) = 0$, the DM receives no information on the comparison. The arrival rate of local signals is normalized by the range of undominated prices $\overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)$. This amounts to assuming that the DM dedicates a fixed amount of attention to the range of prices relevant to the valuation of x regardless of its "size", and ensures that valuations are invariant to affine reparameterizations of the price structure, i.e. denominating certainty equivalents in terms of dollars vs. cents.

This model describes a process where the DM narrows down her valuation of x by querying whether x is worth more or less than different prices in $\mathcal{Z}$, where queries are less likely to be informative when they involve less comparable prices. This leaves the DM with a range of uncertainty over x's valuation that is more likely to include harder-to-compare prices.

Let $V_{\mathcal{Z}}(x)$ denote the DM's valuation. It can be shown that $V_{\mathcal{Z}}(x)$ is distributed according to

$$V_{\mathcal{Z}}(x) = v_{\mathcal{Z}}(x) + \frac{1}{2}\left(\overline{\delta}_{\mathcal{Z}}(x) - \underline{\delta}_{\mathcal{Z}}(x)\right),$$

where $\overline{\delta}_{\mathcal{Z}}(x)$ and $\underline{\delta}_{\mathcal{Z}}(x)$ are independent and are distributed, respectively, according to

$$\underline{F}^x_{\mathcal{Z}}(t) = \begin{cases} 1 - \exp\left(-\int_0^t \lambda^x_{\mathcal{Z}}(v_{\mathcal{Z}}(x) - r)\,dr\right) & t < v_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x) \\ 1 & \text{otherwise} \end{cases}$$

$$\overline{F}^x_{\mathcal{Z}}(t) = \begin{cases} 1 - \exp\left(-\int_0^t \lambda^x_{\mathcal{Z}}(v_{\mathcal{Z}}(x) + r)\,dr\right) & t < \overline{w}_{\mathcal{Z}}(x) - v_{\mathcal{Z}}(x) \\ 1 & \text{otherwise} \end{cases}$$

Importantly, the error $\frac{1}{2}\left(\overline{\delta}_{\mathcal{Z}}(x) - \underline{\delta}_{\mathcal{Z}}(x)\right)$ can be systematically biased as we show later in this section.

Unbounded valuation tasks. In some applications of interest, the valuation task $(x, \mathcal{Z})$ does not contain dominance boundaries. To model these situations, we restrict our attention to valuation tasks where the range of prices is bounded so that $[\underline{w}_{\mathcal{Z}}(x), \overline{w}_{\mathcal{Z}}(x)]$ is an interval, where $\underline{w}_{\mathcal{Z}}(x), \overline{w}_{\mathcal{Z}}(x)$ are defined as above. This is a feature of standard elicitation schemes such as the multiple price list, which restrict the range of prices subjects face by design; in practice, this range is typically specified so that it contains the set of "reasonable" valuations.[7] Here, the signal process can be defined exactly as above.

[7] For instance, when eliciting the valuation of $x = (\$6, 1)$ against $z_w = (\$w, 0.2)$ in a multiple price list, which is not unbounded by dominance from above, McGranaghan et al. (2024b) restrict the range of price list to $[\$6, \$36]$.

Microfoundation of signal structure. This signal structure can be microfounded as a limit case of Shubatt and Yang (2024), who develop and apply a model of choice under imprecise comparisons to study valuations elicited from a multiple price list. In their framework, the DM completes a multiple price list on the basis of noisy pairwise comparisons between the option being valued and the prices in the list, which leads to patterns of systematic noise in valuations governed by the options' comparability to prices. By putting structure on the ease of comparison—an approach we follow below—Shubatt and Yang (2024) show how their model can explain a range of empirical regularities spanning valuation and choice. In Appendix A.1, we show that our signal structure can be derived as a limit case of a version of their model in which the price list becomes arbitrarily fine-grained. In this sense, our model captures the same predictions and underlying psychology as Shubatt and Yang (2024) while providing a more tractable representation.

3.3 Specifying Comparison Difficulty

The inputs to the model are the DM's preference $\succeq$ and the comparability function $\tau(\cdot,\cdot)$. In this paper, we will focus on expected utility as our main specification for $\succeq$, and follow Shubatt and Yang (2024) in specifying $\tau(\cdot,\cdot)$ as being governed by the difficulty of making tradeoffs.

Suppose that $\succeq$ has an expected utility (EU) representation given by the strictly increasing Bernoulli utility function $u : \mathbb{R} \to \mathbb{R}$. Our main specification of $\tau(\cdot,\cdot)$ takes the following form:

Definition 1. *A comparability function τ has a CDF-complexity representation (u, H) if there exists strictly increasing u such that for all $x \neq y$,*

$$\tau(x,y) = H\left(\frac{|EU(x) - EU(y)|}{d_{CDF}(x,y)}\right)$$

for H strictly increasing with $H(0) = 0, H(1) = 1$, where $EU(x) = \sum_w u(w) f_x(w)$ and

$$d_{CDF}(x,y) = \int_0^1 |u(F_x^{-1}(q)) - u(F_y^{-1}(q))|\, dq.$$

Here, comparability is increasing in the *CDF ratio* $r_{CDF}(x,y) \equiv \frac{|EU(x)-EU(y)|}{d_{CDF}(x,y)}$. This ratio captures the severity of tradeoffs in the comparison between x and y. Intuitively, holding fixed the utility difference of x and y, as their payoff distributions become more dissimilar—that is, as $d_{CDF}(x,y)$ grows—the DM faces greater tradeoffs across lottery outcomes, and so the ease of comparison decreases under $r_{CDF}(x,y)$.

This comparability function satisfies several intuitive properties. First, it satisfies *dominance*: $\tau(x,y)$ takes on its maximal value of 1 if and only if x and y have a dominance relationship. It also satisfies *monotonicity*: if $x \succeq y$, then $x' >_{FOSD} x$ implies $\tau(x',y) > \tau(x,y)$ and $y >_{FOSD} y'$ implies $\tau(x,y) < \tau(x,y')$. As such, the measure can be interpreted as capturing the "distance to dominance" in a comparison. The function additionally satisfies *linearity*: $\tau(x,y) = \tau(\lambda x + (1-\lambda)z, \lambda y + (1-\lambda)z)$ for all $\lambda \in (0,1)$, an analog of Independence.

Shubatt and Yang (2024) provide formal foundations for this complexity measure, and show that, when embedded in a model of choice under imprecise comparisons, it can rationalize behavioral regularities such as inverse-S probability weighting in certainty equivalents, preference reversals between valuation and choice, and the instability of probability weighting across valuation formats. Empirically, the measure is highly predictive of the difficulty of binary lottery choice as measured by choice inconsistency and reported subjective uncertainty (Enke and Shubatt, 2023; Shubatt and Yang, 2024).

Parameterizing the model. When deriving numerical predictions from our model, we will use the parameterization $H(r) = 1-(1-r)^{\gamma}$, where $\gamma > 0$ determines the rate at which comparison difficulty increases away from dominance.[8] Under this parameterization, given the Bernoulli utility function u, a standard preference object, our specification of noise is entirely pinned down by the single additional scalar parameter γ.

Relaxing EU preferences. Our specification of tradeoff-driven noise is defined for a DM with EU preferences. In Appendix A.3, we show how this specification can be extended to rank-dependent utility (Kahneman and Tversky, 1979), a prominent theory of non-EU preferences.

It is worth emphasizing that in assuming EU preferences, our goal is not to argue that non-standard risk preferences do not exist or are quantitatively unimportant. Rather, one of our goals will be to investigate the extent to which documented valuation patterns, many of which challenge both expected utility and canonical non-EU risk preferences under non-systematic noise, can be rationalized by EU preferences and our account of tradeoff-driven noise. As discussed in Section 4, we find that while our model can explain first-order features of these patterns, they contain certain features that the model cannot rationalize, which highlights areas in which our assumed theory of preferences (EU) may be inadequate. In Section 5, we build on this idea and develop a framework to speak to which kinds of valuation data are informative about underlying preferences, in the presence of tradeoff-driven noise.

[8] Shubatt and Yang (2024) use an analogous parameterization in their specification of comparison precision.

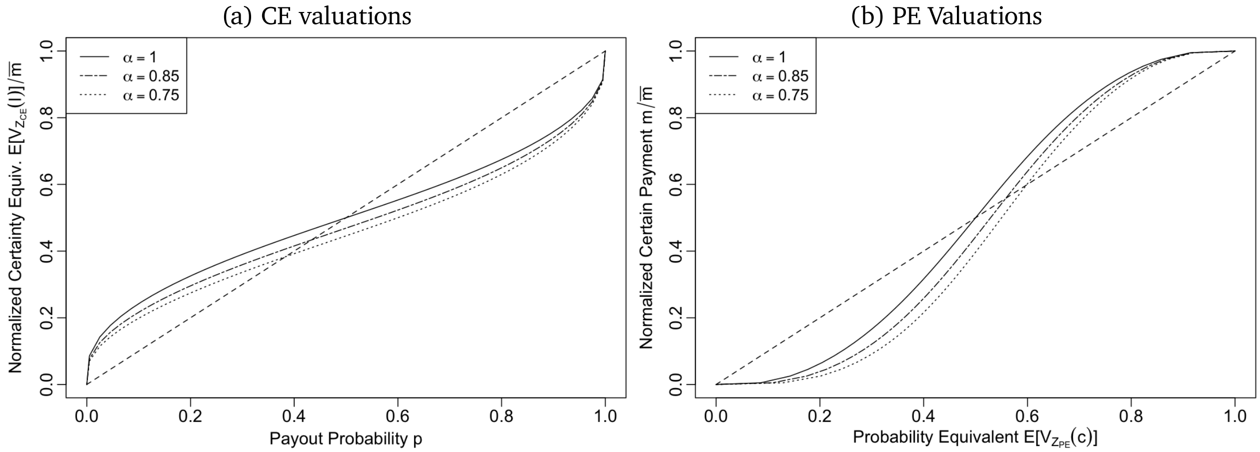

Figure 3: Predicted certainty equivalents $\mathbb{E}[V_{\mathcal{Z}_{CE}}(\ell)]$ (panel a) and probability equivalents $\mathbb{E}[V_{\mathcal{Z}_{PE}}(c)]$ (panel b). Preferences are given by $u(w) = w^{\alpha}$, and the comparability function τ has a CDF-complexity representation (u, H) with $H(r) = 1 - (1-r)^{\gamma}$, for $\gamma = 2.5$.

3.4 Noise and Systematic Compression

A key feature of our model of noise is that it predicts systematic noise in valuation in the form of a compression effect: when tradeoffs make x hard to compare to a range of prices, its valuation is pulled toward the center of that range. Intuitively, this is because the DM's posterior range of uncertainty over x's valuation is more likely to contain prices that are difficult to compare. To illustrate, we consider two standard valuation paradigms from the literature.

Certainty equivalents. The DM values a binary lottery $\ell = (\overline{m}; p)$ against money $\mathcal{Z}_{CE} = \{(w; 1)_{w \in \mathbb{R}}\}$. As the lottery's payoff probability p departs from the boundaries of 0 and 1, risk/return tradeoffs make ℓ difficult to compare to an increasingly wider range of prices in $[0, \overline{m}]$. This compresses its valuation toward the center of that range, as illustrated in Figure 3a, which plots average normalized valuations $\mathbb{E}[V_{\mathcal{Z}_{CE}}(\ell)]/\overline{m}$ predicted by our model as a function of p. Notice two key features of valuation noise. First, compression leads to systematic distortions in levels: low probability lotteries are overvalued and high probability lotteries are undervalued. Second, compression generates attenuation (Enke et al., 2024): for interior p, valuations are insufficiently sensitive to variation in p. As such, our model of noise can produce apparent inverse-S probability weighting in valuations without appealing to non-standard preferences, and coheres with prior work providing complexity-based foundations for probability weighting (Enke and Graeber, 2023; Shubatt and Yang, 2024; Frydman and Jin, 2025; Oprea, 2024).

Probability equivalents. The DM values a certain payment $c = (m; 1)$ against a yardstick lottery $\mathcal{Z}_{PE} = \{(\overline{m}; w)\}_{w \in [0,1]}$. Here, the difficulty of making risk/return tradeoffs again leads to systematic compression. This is illustrated in Figure 3b, which plots $\mathbb{E}[V_{\mathcal{Z}_{PE}}(c)]$ (x-axis) as a function of $m/\overline{m}$ (y-axis): low certain payments are overvalued and high certain payments are undervalued, leading to a *reversal* of inverse-S probability weighting (Feldman and Ferraro, 2024; Shubatt and Yang, 2024). This highlights a key feature of our model: since noise depends on the tradeoffs in comparing the option to the numeraire, it can manifest differently across valuation tasks.

Relationship to naive compression. Compression effects in valuations can be explained by other accounts of noise, such as a model in which valuations are mechanically biased toward the "middle" of the price list, or a model in which a proportion of subjects state valuations uniformly at random in a fixed range. Unlike these accounts of naive compression, however, our model makes predictions about structure of compression effects. For instance, let $x = (\$30; 0.79)$. While we might expect compression in the valuation of x against $\{(\$w; 1)\}_{w\in\mathbb{R}}$, we should expect little noise or compression in the valuation of x against $\{(\$w; 0.80)\}_{w\in\mathbb{R}}$, as here x has a near-dominance relationship with most of the prices. Likewise, while we might expect compression in both the valuation of $x = (\$30; 0.79)$ and $y = (\$5; 0.79)$ against $\{(\$w; 1)\}_{w\in\mathbb{R}}$, we should not expect the valuations of x and y to be compressed toward the same "intermediate" value. By putting structure on the likely range of uncertain valuations over which compression occurs, our model naturally captures these regularities. As the next section illustrates, this added structure will be key to rationalizing valuation inconsistencies we document, as well as a set of patterns from the extant literature.

4 Explaining Valuation Patterns

We show how our model of EU preferences and tradeoff-driven noise can rationalize a range of documented patterns that are difficult to rationalize under prevailing theories of risk preferences under standard assumptions on noise. We begin with our own experimental results.

4.1 Motivating Evidence

The experiment in Section 2 considered the valuation of the two lotteries $\ell_R = (30; 0.2 \cdot p)$, $\ell_S = (30 \cdot p; 0.2)$ against four price structures:

$$\mathcal{Z}_{M\text{-}a} = \{(w; 0.2)\}_{w\in\mathbb{R}} \qquad \mathcal{Z}_{M\text{-}b} = \{(w; 0.5)\}_{w\in\mathbb{R}}$$
$$\mathcal{Z}_{P\text{-}a} = \{(30; w)\}_{w\in[0,1]} \qquad \mathcal{Z}_{P\text{-}b} = \{(40; w)\}_{w\in[0,1]}$$

Our noise specification rationalizes the inconsistencies we document as a result of tradeoff-driven compression.

First, consider M-a valuations. Here, the difficulty of comparing ℓ_R to the prices in $\mathcal{Z}_{M\text{-}a}$ results in compression effects similar to those discussed in Section 3.4, which cause ℓ_R to be overvalued for low p and undervalued for high p. On the other hand, ℓ_S is trivial to compare to prices, resulting in no distortions. As the numerical results in Figure 4a illustrate, for a range of true underlying risk preferences, this results in valuations that are risk-seeking for low p, and risk-averse for high p.

Now consider P-a valuations. Here, the situation is exactly reversed relative to M-a valuations. ℓ_R is now trivial to compare to prices, whereas tradeoffs generate compression effects in the valuation of ℓ_S, which as a result is overvalued (undervalued) for low (high) p. As shown in Figure 4b, valuations here are risk-averse for low p and risk-seeking for high p.

Now consider M-b valuations. Here, the DM faces tradeoffs in comparing both ℓ_R and ℓ_S to the prices in $\mathcal{Z}_{M\text{-}b}$. Here, as p decreases, the resulting compression causes valuations of ℓ_R to be insufficiently sensitive to the resulting decrease in its payoff probability, whereas the valuations

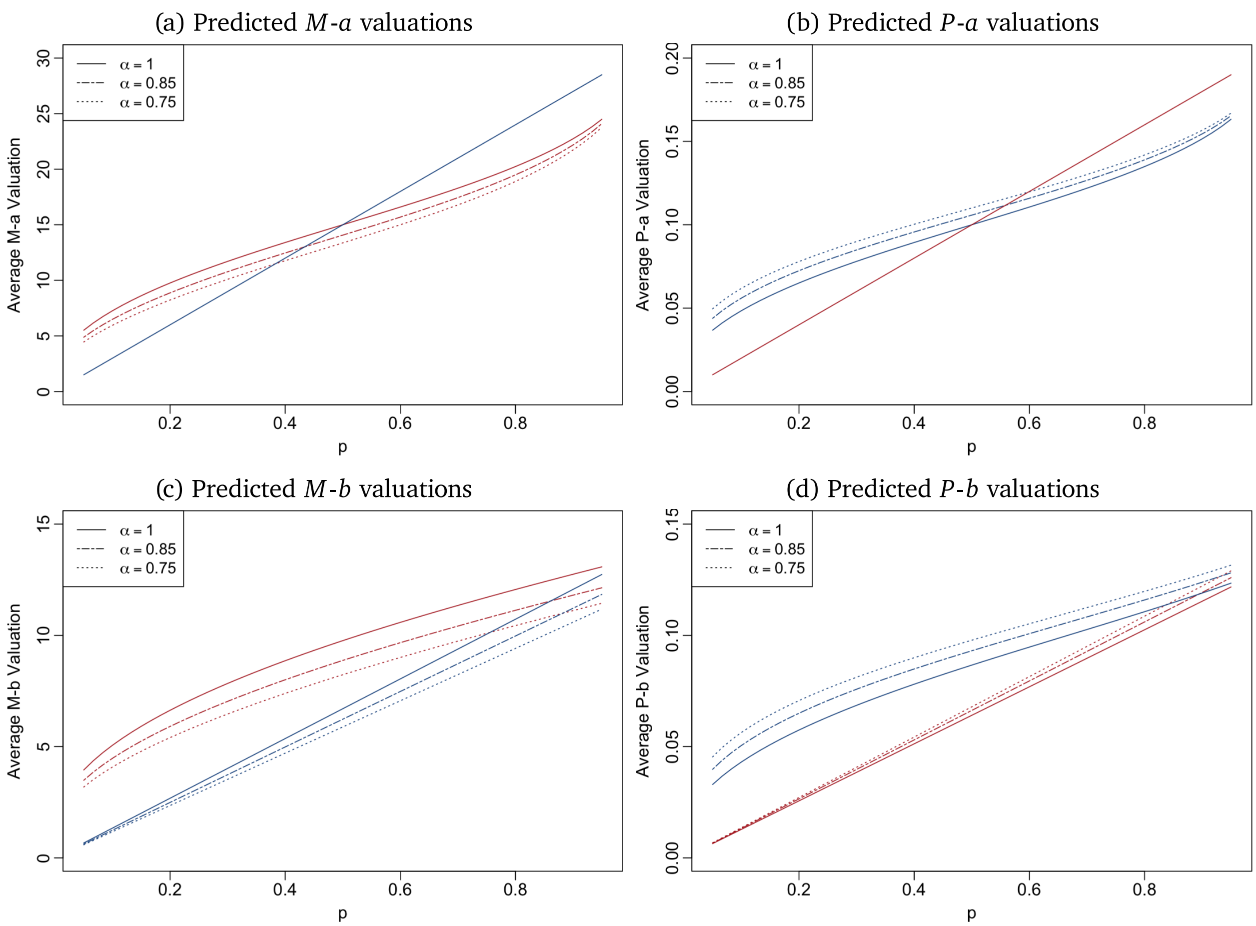

Figure 4: Predicted valuations of ℓ_R (in red) and ℓ_S (in blue) against $\mathcal{Z}_{M\text{-}a}$, $\mathcal{Z}_{P\text{-}a}$, $\mathcal{Z}_{M\text{-}b}$, $\mathcal{Z}_{P\text{-}b}$. Preferences are given by $u(w) = w^{\alpha}$, and the comparability function τ has a CDF-complexity representation (u, H) with $H(r) = 1-(1-r)^{\gamma}$, for $\gamma = 2.5$.

of ℓ_S remain more sensitive to the resulting decrease in its payoff amount, which governs the range of prices that are undominated with respect to ℓ_S.[9] As a result, valuations increasingly favor ℓ_R as p decreases, which as the numerical results in Figure 4c show, can cause valuations to be risk-seeking for the range of p we consider in our experiment.

Finally, consider P-b valuations. Again, the DM faces tradeoffs in comparing both ℓ_R and ℓ_S to prices. In this case, however, compression instead causes valuations to be insensitive to variation in payoffs, rather than probabilities. This generates the opposite pattern as in the case of M-b valuations: valuations increasingly favor ℓ_S over ℓ_R as p decreases, which can cause valuations to be risk-averse over the relevant range of p, as Figure 4d shows.

Figure 4 shows that our model of noise can explain the inconsistencies we observe in our experiment. In Appendix A.7, we provide formal and numerical results demonstrating that our model is restrictive, in that it rules out the counterfactual inconsistencies that we do not observe in our experiment.

4.2 Compounded Valuation Tasks

Consider the following valuation tasks:

$$\textit{CEs:} \quad \ell(p) = (\overline{m}, p) \text{ valued against } \mathcal{Z} = \{(w,1)\}_{w\in\mathbb{R}}$$
$$\textit{Compounded CEs:} \quad \ell_r(p) = (\overline{m}, rp) \text{ valued against } \mathcal{Z}_r = \{(w,r)\}_{w\in\mathbb{R}}$$

where the compounded CE task is obtained by scaling down the payoff probabilities of all the options in the CE task by a common factor $r \in [0,1]$.

McGranaghan et al. (2024b) elicit valuations using both tasks as a test of common ratio preferences. They find that certainty equivalents exhibit the standard pattern of apparent inverse-S probability weighting: valuations are risk-seeking for low probabilities and risk-averse for high probabilities; at the same time, they find no aggregate effect of compounding: in their data, $E[V_{\mathcal{Z}}(\ell(p))] \approx E[V_{\mathcal{Z}_r}(\ell_r(p))]$ for all ℓ, r. That is, both CEs and the compounded version of the task exhibit the *same* inverse-S pattern.

As McGranaghan et al. (2024b) note, their results are puzzling from the perspective of prevailing models of risk preferences. The inverse-S pattern in certainty equivalents is difficult to reconcile with standard parameterizations of expected utility, and indeed, is often taken as motivating evidence against expected utility in favor of models of non-EU preferences.[10] At the same time, the most commonly cited explanation for this pattern—non-linear probability

[9]This can be seen as a formalization of the "compatibility effect" discussed in (Tversky et al., 1990), who posited that judgments are more responsive to variation in features that are more congruent with the response scale.

[10]Caveat: While EU preferences can explain this inverse-S pattern through specifications of utility curvature...

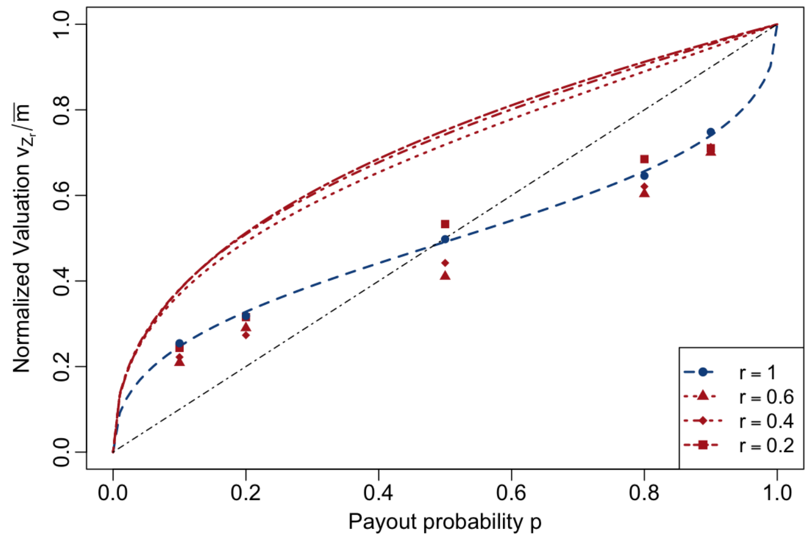

Figure 5: Predicted compounded certainty equivalents under CPT. Lines plot predicted valuations $v^{CPT}_{\mathcal{Z}_r}(\ell_r(p))$ of a CPT model estimated on the certainty equivalents data from McGranaghan et al. (2024b), under the parameterization $u(x) = x$, $\pi(q) = \chi q^\nu/(\chi q^\nu + (1-q)^\nu)$. Points plot observed valuations $\mathbb{E}[V_{\mathcal{Z}_r}(\ell_r(p))]$ in McGranaghan et al. (2024b).

weighting in cumulative prospect theory (Tversky and Kahneman, 1992)—is inconsistent with the invariance of CEs to compounding.

To see this, consider the predicted valuations of cumulative prospect theory (CPT), in which the utility of a binary lottery $(m; q)$ is given by $\pi(q)u(w)$ for a strictly increasing weighting function $\pi : [0, 1] \to [0, 1]$ satisfying $\pi(0) = 0$, $\pi(1) = 1$ and strictly increasing u. Under CPT, the DM's valuation of $l_r(p)$ is given by

$$v^{CPT}_{\mathcal{Z}_r}(l_r(p)) \equiv u^{-1}\left(\frac{\pi(pr)}{\pi(r)}u(\overline{m})\right).$$

This implies that valuations will be invariant to compounding r if and only if π is linear, i.e. the DM has EU preferences. Figure 5 illustrates: calibrating CPT to the certainty equivalents data from McGranaghan et al. (2024b) (dashed blue curve), CPT predicts a markedly different pattern in compounded valuations (red dashed/dotted curves). In contrast, McGranaghan et al. (2024b) find that in the aggregate, compounded valuations exhibit virtually the same pattern.

Our model can rationalize these patterns as a result of tradeoff-driven noise, under expected utility preferences. Intuitively, our model generates apparent inverse-S probability weighting as a consequence of the difficulty of trading off probabilities and payouts in certainty equivalent tasks, a difficulty that remains even when the prospects in the task are compounded. Indeed, our model predicts no effect of compounding:

Proposition 1. *Suppose $\succeq$ has an EU representation u and $\tau(\cdot,\cdot)$ has a CDF-complexity represen-*

tation (u,H). For $\ell_r(p)$ and $\mathcal{Z}_r$ as defined above, for all p, $\mathbb{E}[V_{\mathcal{Z}_r}(\ell_r(p))]$ is constant in $r \in (0,1]$.

This is because under CDF-complexity, the comparability between $l_r(p)$ and each price in $\mathcal{Z}_r$ is invariant to r. At the same time, however, compression effects in our model can generate apparent inverse-S probability weighting, as Section 3.4 illustrates. Our model can therefore rationalize the simultaneous presence of apparent non-linear probability-weighting in certainty equivalents, as well as its invariance to compounding.

Importantly, while McGranaghan et al. (2024b) find no aggregate effect of compounding, they find stable responses to compounding on the individual level, which they interpret as reflecting non-EU preferences; McGranaghan et al. (2024a) also find evidence of systematic responses to compounding in a variation of their paired valuation paradigm consistent with an aggregate common ratio preference. Our model of tradeoff-driven noise rules out these patterns under EU preferences, suggesting that they may be informative about underlying non-EU preferences. We return to this point in Section 5.

4.3 Equalizing Reductions and Rank-Independence

Bernheim and Sprenger (2020) introduce the *equalizing reductions* (ER) paradigm to test a key implication of cumulative prospect theory: that if preferences exhibit non-linear probability weighting, they must also exhibit rank-dependence. Under CPT, the utility of a lottery $x = (a,b,c;p,q,1-p-q)$ is given by

$$U_{CPT}(x) = w_a u(a) + w_b u(b) + w_c u(c)$$

CPT predicts that if the decision weights (w_a, w_b, w_c) exhibit non-linear probability weighting, then they must depend on the ranks of the payoffs (a, b, c). The equalizing reductions paradigm is designed to test this prediction.

In the equalizing reductions task, the subject reports the reduction in payoff c required to offset a fixed increase in the payoff b. That is, for $\overline{w} > 0$, the DM values

$$x = (a,b,c;p,q,1-p-q) \text{ against } \mathcal{Z}_{ER} = \{(a, b+m, c-w; p, q, 1-p-q)\}_{w\in[0,\overline{w}]}.$$

For small m, the normalized equalizing reduction $v_{\mathcal{Z}_{ER}}(x)/m$ identifies the relative decision weights w_b/w_c up to a scaling constant. Bernheim and Sprenger (2020) exploit this observation in two ways. First, by varying p and q, they test for non-linear probability weighting: for m small, $v_{\mathcal{Z}_{ER}}(x)/m$ is linear in $q/(1-p-q)$ under EU, and so any non-linearities in the relationship are indicative of non-linear probability weighting. Second, by varying a, they change the rank ordering of payoffs and thereby test the rank-dependence implication of CPT, which predicts that $v_{\mathcal{Z}_{ER}}(x)/m$ should change discontinuously with the rank ordering of payoffs so long as there is non-linear probability weighting.

They document two key findings: 1) the observed equalizing reductions $\mathbb{E}[V_{\mathcal{Z}_{ER}}(x)]/m$ exhibit diminishing sensitivity with respect to $q/(1-p-q)$, indicative of pronounced non-linear probability weighting; and 2) $\mathbb{E}[V_{\mathcal{Z}_{ER}}(x)]/m$ is essentially unresponsive to variation in the rank ordering of payoffs triggered by variation in a, which strongly rejects rank-dependence. Taken at face value, these valuation behaviors are inconsistent with both expected utility, which cannot explain 1), and cumulative prospect theory, which cannot explain both 1) and 2).

Our model of EU preferences and tradeoff-driven noise can jointly explain these patterns. First, our model delivers the sharp prediction that equalizing reductions are rank-independent, and in particular are invariant to variation in the common payoff a:

Proposition 2. *Suppose $\succeq$ has an EU representation u and $\tau(\cdot,\cdot)$ has a CDF-complexity representation (u,H). For $x=(a,b,c;p,q,1-p-q)$ and $\mathcal{Z}_{ER}=\{(a,b+m,c-w;p,q,1-p-q)\}_{w\in[0,\overline{w}]}$ for any $\overline{w}>0$, $\mathbb{E}[V_{\mathcal{Z}_{ER}}(x)]$ is constant in a.*

The intuition is as follows: under CDF complexity, the comparability function τ satisfies a cancellation property: the ease of comparison between two lotteries is invariant to common consequences that the lotteries share. As a result, varying the common consequence a changes neither the true equalizing $v_{\mathcal{Z}_{ER}}(x)$ reduction under EU preferences nor the ease of comparing x to the prices in $\mathcal{Z}_{ER}$, and so valuations exhibit rank-independence.

At the same time, our specification of noise can generate apparent non-linear probability weighting in how equalizing reductions vary with $q/(1-p-q)$, even if underlying preferences are EU. Intuitively, as $q/(1-p-q)$ increases away from 0, where the valuation problem is trivial, the DM faces increasingly difficult tradeoffs in comparing x to prices in $\mathcal{Z}_{ER}$, expanding the range of uncertainty of x's valuation; this in turn leads to greater compression effects and so produces diminishing sensitivity of equalizing reductions to $q/(1-p-q)$.

As the calibration exercise in Figure 6 shows, this force can quantitatively explain the observed diminishing sensitivity of equalizing reductions to probabilities observed in Bernheim and Sprenger (2020). Here, we calibrate the parameters of our model against certainty equivalents data from Bernheim and Sprenger (2020), in which $\ell=(\overline{m};p)$ is valued against $\mathcal{Z}_{CE}=(w;1)_{w\in\mathbb{R}}$, and assess the resulting out-of-sample predictions of our model to the equalizing reductions data that the authors collect in the same sample. Panel a) plots the in-sample predictions of our model to certainty equivalents, and Panel b) plots the out-of-sample predictions for equalizing reductions; here, we see that our model predicts the diminishing sensitivity of equalizing reductions to payoff probabilities that Bernheim and Sprenger (2020) document.

In summary, we show that a standard account of risk preferences can account for puzzling

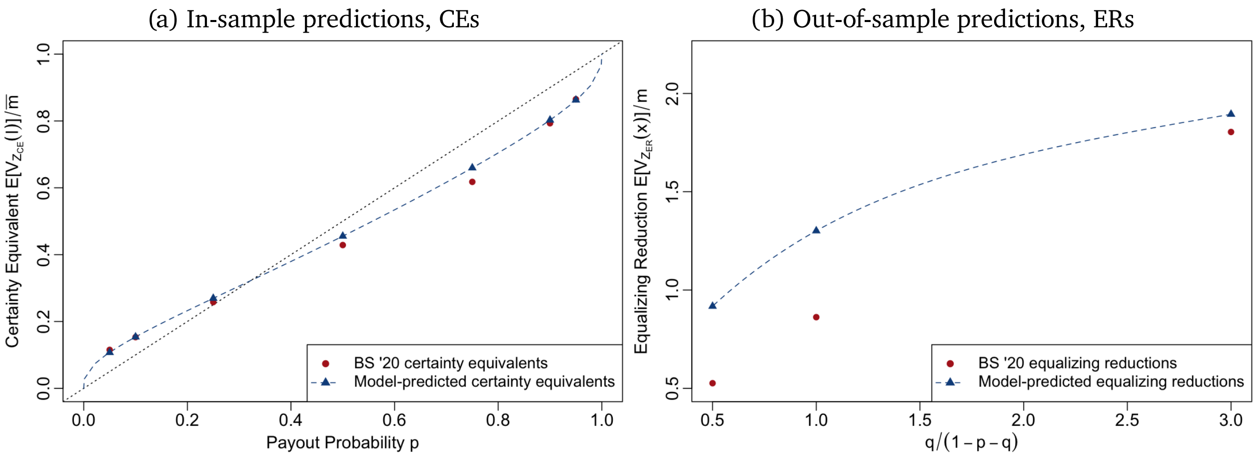

Figure 6: In-sample model predictions for certainty equivalents (Panel a) and out-of-sample predictions for equalizing reductions (Panel b) in the data of Bernheim and Sprenger (2020). For certainty equivalents, $\ell = (25; p)$ is valued against $\mathcal{Z}_{CE} = \{(w;1)\}_{w\in\mathbb{R}}$. For equalizing reductions, $x = (a, 24, 18; p, q, 1-p-q)$ is valued against $\mathcal{Z}_{ER} = \{(a, 24+m, 18-w)\}_{w\in[0,16]}$, pooling valuations across $a \in \{23, 21, 19\}$. Preferences are given by $u(w) = w^{\alpha}$, and τ has a CDF-complexity representation (u, H) with $H(r) = 1-(1-r)^{\gamma}$. Parameter values $(\alpha, \gamma) = (0.85, 5.5)$ are calibrated to the certainty equivalents data of Bernheim and Sprenger (2020).

patterns in equalizing reductions—rank independence in response to variation in common payoffs and non-linear probability weighting in response to variation in probabilities—under our specification of systematic, complexity-driven noise.[11]

4.4 Uncertainty Equivalents

Andreoni and Sprenger (2011) devise a test of independence using the uncertainty equivalents (UE) valuation paradigm, where given positive payouts m_L, m_H with $m_L < m_H$, the lottery

$$\ell(p) = (p, 1-p; m_L, m_H) \text{ is valued against } \mathcal{Z} = \{(m_H; w)\}_{w\in[0,1]}.$$

Here, the DM provides the probability q that makes the binary lottery $(q; w_H)$ indifferent to ℓ_p.

Andreoni and Sprenger (2011) show that eliciting uncertainty equivalents across the range of $p \in [0,1]$, allows for a test of the independence axiom: under EU preferences, q should be linear in p, whereas a class of non-EU risk preferences, such as forms of non-linear probability weighting and disappointment aversion (Bell, 1985; Loomes and Sugden, 1986; Gul,

[11]Bernheim and Sprenger (2020) also document a form of complexity aversion in certainty equivalents, where the valuation of a lottery decreases sharply when a payoff outcome is split into two outcomes, e.g. 60% of receiving \$20 is split into 30% of receiving $\$20+\epsilon$ and 30% of receiving $\$20-\epsilon$. As they discuss, this complexity aversion implies dominance violations in certainty equivalents. As valuations in our model respect a form of dominance (see Appendix A.2), our model cannot easily rationalize these findings, consistent with the interpretation in Bernheim and Sprenger (2020) that these patterns reflect an underlying non-EU preference for simplicity.

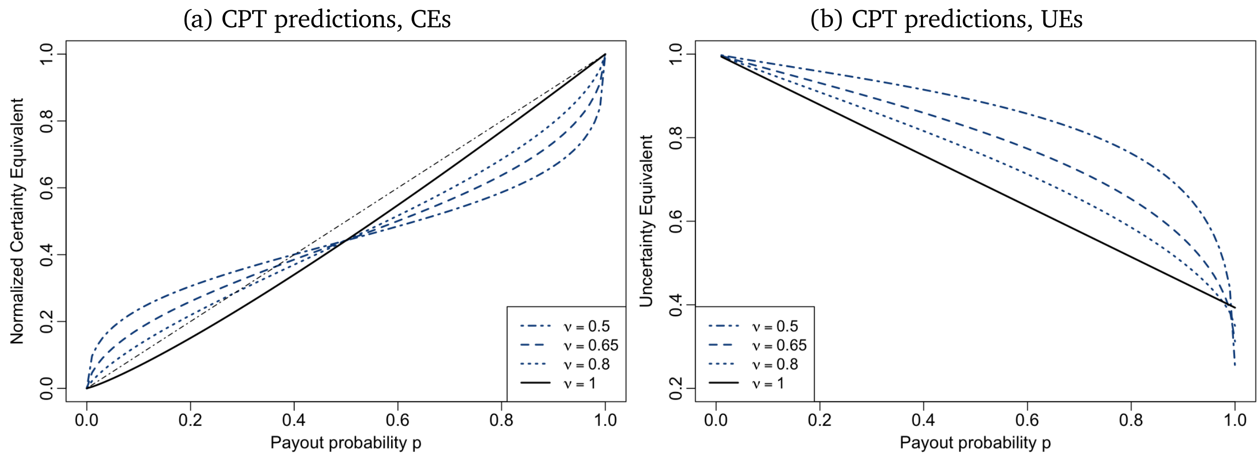

Figure 7: CPT-predicted certainty equivalents and uncertainty equivalents. CPT is parameterized by $u(x) = x^\alpha$, $\pi(q) = q^\nu/(q^\nu + (1-q)^\nu)$, with $\alpha = 0.85$, $\nu \in \{0.4, 0.55, 0.7, 1\}$. For certainty equivalents, $\ell_{CE}(p) = (30; p)$ is valued against $\mathcal{Z}_{CE} = \{(w; 1)\}_{w\in\mathbb{R}}$. For uncertainty equivalents, $\ell_{UE}(p) = (10, 30; p, 1-p)$ is valued against $\mathcal{Z}_{UE} = \{(30; w)\}_{w\in[0,1]}$.

1991), predict a non-linear relationship between q and p. To illustrate, Figure 7 plots the uncertainty equivalents predicted by various parameterizations of CPT, alongside the corresponding predictions for standard certainty equivalents. The figure makes clear that these predictions are linked: parameterizations that generate the classic inverse-S pattern in CEs also imply a markedly non-linear relationship between valuations and p in UEs. Thus, under CPT, the same probability-weighting mechanism used to explain inverse-S curvature in certainty equivalents necessarily predicts substantial curvature in uncertainty equivalents.

In contrast, Andreoni and Sprenger (2011) find that along nearly the full range of $p \in [0, 1]$, uncertainty equivalent valuations are essentially linear in p, despite the fact that the certainty equivalents in the same experiment exhibit the standard inverse-S weighting pattern. Taken at face value, these results are puzzling both from the perspective of expected utility and prevailing non-EU preferences.

Our specification of tradeoff-driven noise and EU preferences can account for these patterns. While our model predicts apparent inverse-S weighting in certainty equivalents due to noise-driven compression, it simultaneously makes the sharp prediction that uncertainty equivalents are linear in p:

Proposition 3. *Suppose $\succeq$ has an EU representation u and $\tau(\cdot,\cdot)$ has a CDF-complexity representation (u, H). For $\ell(p)$ and $\mathcal{Z}$ as defined above, $1 - \mathbb{E}[V_{\mathcal{Z}}(\ell(p))]$ are linear in p.*

Intuitively, even though tradeoff-driven compression systematically distorts the uncertainty equivalents of $\ell(p)$ for any given p, these distortions scale proportionally with p under CDF complexity so as to maintain linearity. Our model therefore implies that a key feature of uncertainty equivalents—linearity under EU preferences—is preserved even in the presence of tradeoff-driven noise.

While Andreoni and Sprenger (2011) find that probability equivalents are linear in p along nearly the entire unit interval, they do document one instance of non-linearity: valuations of l_p exhibit a slight upward deviation from this linear relationship for p close to 1. They interpret this "certainty premium" pattern as evidence for a systematic preference for certainty (Neilson, 1992; Schmidt, 1998; Diecidue et al., 2004), in which the DM attaches a premium to near-riskless lotteries. Our theory of noise rules out these non-linearities under EU preferences, and is thus consistent with the conclusion that this feature of the data reflects underlying non-EU preferences. We return to this point in Section 5.

4.5 Boundary Effects in Valuations

Kontek (2018) documents "boundary effects" in the certainty equivalents of lotteries taking the form $\ell = (m_1, m_2, m_3; p_1, p_2, p_3)$, where $m_1 < m_2 < m_3$, visualized in Figure 8: holding fixed p_3, certainty equivalents exhibit a sharp drop when the probability of the low payout p_1 increases above 0; conversely, holding fixed p_1, certainty equivalents exhibit a sharp rise when the probability of the high payout p_3 increases above 0.

These boundary effects imply that indifference curves on the Marschak-Machina simplex exhibit sharp jumps in slope at the boundaries corresponding to $p_1 = 0$ and $p_3 = 0$ toward the origin, where $p_1 = p_3 = 0$. These jumps are inconsistent with EU, which predicts linear and parallel indifference curves on the Marschak-Machina simplex, as well as CPT, which predicts smooth indifference curves. Kontek (2018) shows that while these canonical models of risk

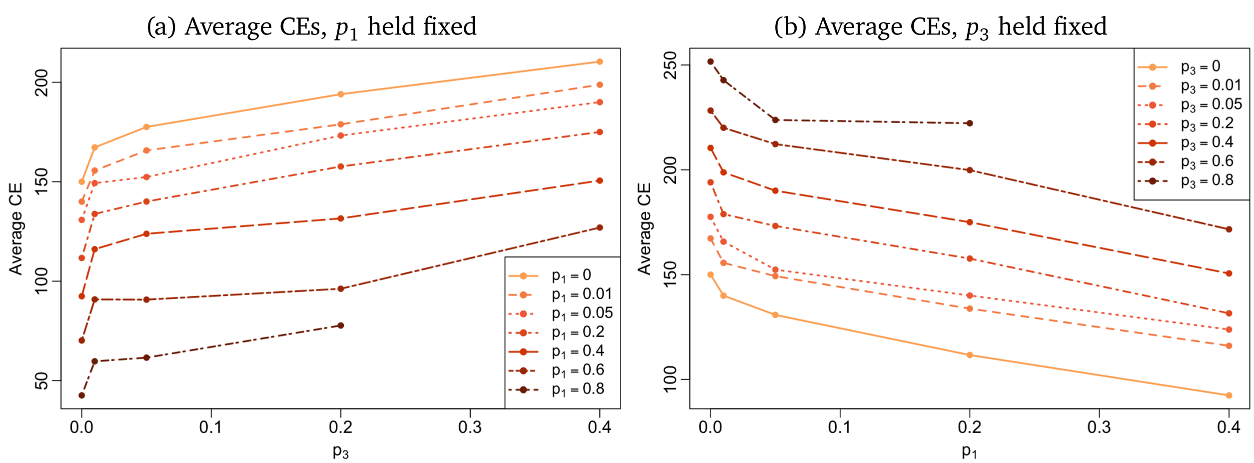

Figure 8: Average certainty equivalents for the lottery $\ell = (m_1, m_2, m_3; p_1, p_2, p_3)$ from Kontek (2018). Here $m_1 = 0, m_2 = 150, m_3 = 300$ in terms of Polish złoty.

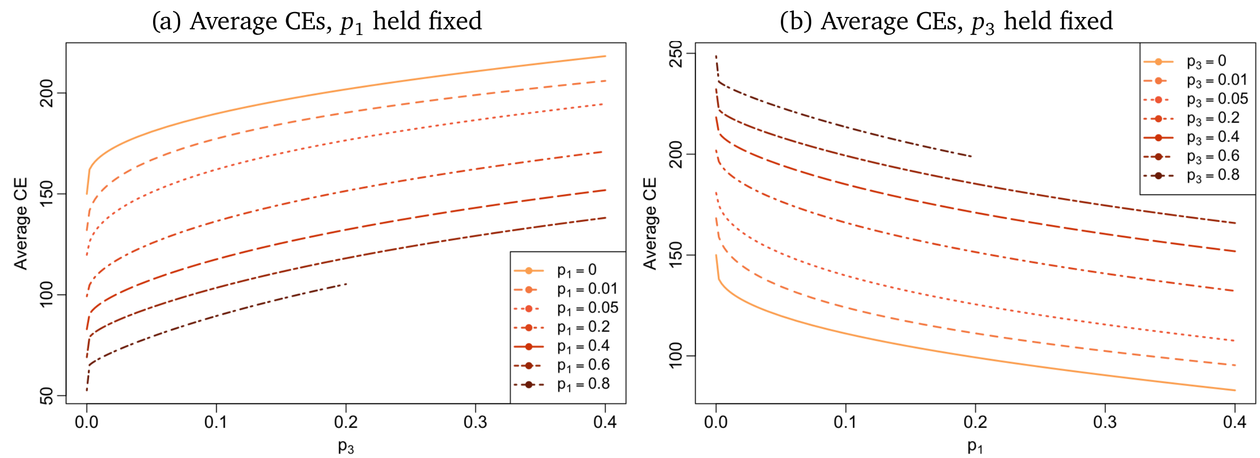

Figure 9: Model-predicted average certainty equivalents $\mathbb{E}[V_{\mathcal{Z}}(\ell)]$ for $\ell = (m_1, m_2, m_3; p_1, p_2, p_3)$ valued against $\mathcal{Z} = \{(w, 1)\}_{w \in \mathbb{R}}$, where $m_1 = 0, m_2 = 150, m_3 = 300$. Preferences are given by $u(w) = w^{\alpha}$, and τ has a CDF-complexity representation (u, H) with $H(r) = 1 - (1-r)^{\gamma}$, with $(\alpha, \gamma) = (1, 2)$.

preferences cannot explain these boundary effects, they can be rationalized by a model of range-dependent risk preferences (Kontek and Lewandowski, 2018).

Our model of noise provides an alternative explanation for these boundary effects. To see why, consider the lotteries

$$x = (m_1, m_2, m_3; 0, 0.8, 0.2) \qquad y = (m_1, m_2, m_3; 0.2, 0.8, 0)$$
$$x' = (m_1, m_2, m_3; 0.01, 0.79, 0.2) \qquad y' = (m_1, m_2, m_3; 0.2, 0.79, 0.01).$$

First, consider the valuation of x, x' in terms of certainty equivalents, i.e. $\mathcal{Z} = \{(w; 1)\}_{w \in \mathbb{R}}$. In the valuation of x, the range of undominated prices is given by $[m_2, m_3]$, whereas in the valuation of x' this range extends downward to $[m_1, m_3]$. As such, the compression effects in our model, which pull valuations to the middle of the range of uncertainty, will penalize the valuation of x' relative to x. The analogous logic predicts a boost in valuations moving from y to y', as here the range of undominated prices extends upward from $[m_1, m_2]$ to $[m_1, m_3]$.

Figure 9 shows that our parameterized model can generate boundary effects that qualitatively match the patterns documented in Kontek (2018).

5 Complexity-Robust Valuation Designs

In this section, we use our model to derive implications for what can be learned about preferences from valuation data in the presence of complexity-driven noise. The key implication of our model here is that while the *levels* of valuations are in general distorted by systematic noise, and should therefore be interpreted with caution, *differences* between valuations can be informative about underlying preferences. Below, we characterize which of these valuation tests are robust to our specification of noise.

5.1 Valuation Paradigms

Given two valuation tasks $(x, \mathcal{Z})$, $(x', \mathcal{Z}')$ and an affine function φ, suppose the analyst is interested in testing whether underlying preferences satisfy the relationship $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$. We will refer to $((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$ as a *valuation paradigm*. Below, we motivate this framework with examples of valuation paradigms from the literature.

Example 2. (Examples of Valuation Paradigms).

- *CE Risk Attitudes.* $x = (\overline{m}; p)$, $x' = (\overline{m} \cdot p; 1)$, $\mathcal{Z} = \mathcal{Z}' = \{(w; 1)\}_{w \in \mathbb{R}}$, $\varphi(v) = v$. This corresponds to the common practice of using certainty equivalents to assess whether the DM is risk averse/risk tolerant, by testing whether the certainty equivalent of a lottery is less than/greater than the lottery's expected value.

- *Paired CRP Test.* $x = (\overline{m}; p)$, $x' = (\overline{m}; rp)$, $\mathcal{Z} = \{(w; 1)\}_{w \in \mathbb{R}}$, $\mathcal{Z}' = \{(w; r)\}_{w \in \mathbb{R}}$, $\varphi(v) = v$. This corresponds to the test for common ratio preferences developed in McGranaghan et al. (2024b). Here, $v_{\mathcal{Z}}(x) > v_{\mathcal{Z}'}(x')$ is indicative of a common ratio preference.

- *Equalizing Reductions: Test of Rank Dependence.*

$$x = (30, 24, 18; p, q, 1-p-q), \quad x' = (21, 24, 18; p, q, 1-p-q)$$
$$\mathcal{Z} = \{(30, 24+m, 18-w; p, q, 1-p-q)\}_{w \in [0, \overline{w}]}$$
$$\mathcal{Z}' = \{(21, 24+m, 18-w; p, q, 1-p-q)\}_{w \in [0, \overline{w}]}$$
$$\varphi(v) = v$$

  This is the test of rank dependence from Bernheim and Sprenger (2020), where the goal is to test whether a change in the rank order of payoffs induced by a variation in the first payoff affects the equalizing reductions between the other two payoffs.

- *Equalizing Reductions: Test of Probability Weighting.*

$$x = (30, 24, 18; p, q, 1-p-q), \quad x' = (30, 24, 18; p', q', 1-p'-q')$$
$$\mathcal{Z} = \{(30, 24+m, 18-w; p, q, 1-p-q)\}_{w \in [0, \overline{w}]}$$
$$\mathcal{Z}' = \{(30, 24+m, 18-w; p', q', 1-p'-q')\}_{w \in [0, \overline{w}]}$$
$$\varphi(v) = \frac{q}{1-p-q} \cdot \frac{1-p'-q'}{q'} \cdot v$$

This is the equalizing reductions test for non-linear probability weighting from Bernheim and Sprenger (2020). The goal is to test whether equalizing reduction w is linear in the relative probabilities $\frac{q}{1-p-q}$. Assuming that the equalizing reduction is 0 when $q = 0$, linearity holds only if $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$.

- *Uncertainty Equivalents Test.* For $0 < m_L < m_H$:

$$x = (m_L, m_H; p, 1-p), \quad x' = (m_L, m_H; p', 1-p')$$
$$\mathcal{Z} = \mathcal{Z}' = \{(m_H; w)\}_{w\in[0,1]}$$
$$\varphi(v) = 1 - \tfrac{p}{p'} \cdot (1-v)$$

  This is the uncertainty equivalents test from Andreoni and Sprenger (2011), where the goal is to test whether uncertainty equivalents of $(m_L, m_H; p, 1-p)$ is linear in p; violations of linearity are indicative of violations of Independence. Assuming that $q = 1$ when $p = 0$, linearity holds only if $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$.

We are interested in whether in the face of noise, a given valuation paradigm provides a valid test of the underlying preference relationship of interest.

Definition 2. *Say that a valuation paradigm $((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$ is valid if whenever $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$, $\mathbb{E}[V_{\mathcal{Z}}(x)] = \mathbb{E}[\varphi(V_{\mathcal{Z}'}(x'))]$.*

In words, a valuation paradigm is valid if whenever the relationship of interest holds for the underlying valuations—that is, if we "assume the null hypothesis"—a comparison of means of the observed valuations provides a valid test of the hypothesis.

Two remarks on this notion of validity are in order. First, note that our definition of validity has a flavor of a frequentist test, in the following sense. Suppose that a paradigm is valid, and the analyst observes the data $\mathbb{E}[V_{\mathcal{Z}}(x)] = \mathbb{E}[\varphi(V_{\mathcal{Z}'}(x'))]$; it does *not* follow that the null hypothesis $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$ holds. Instead, the analyst can only conclude that the data "fail to reject" the null hypothesis. On the other hand, if $\mathbb{E}[V_{\mathcal{Z}}(x)] \neq \mathbb{E}[\varphi(V_{\mathcal{Z}'}(x'))]$, then it cannot be the case that the null hypothesis holds; the analyst can reject the null hypothesis. Second, our validity notion corresponds to a test of valuation means. In Appendix A.4, we show that our results straightforwardly extend to a validity notion that corresponds to a sign test.

5.2 Complexity-Robust Paradigms

The following proposition gives a sufficient criterion for validity of a valuation paradigm.

Proposition 4. *Consider a valuation paradigm $\mathcal{P} = ((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$, where $\varphi(\underline{w}_{\mathcal{Z}'}(x')) = \underline{w}_{\mathcal{Z}}(x)$ and $\varphi(\overline{w}_{\mathcal{Z}'}(x')) = \overline{w}_{\mathcal{Z}}(x)$. Denoting $\mathcal{Z} = \{z_w\}_{w\in\mathcal{W}}$ and $\mathcal{Z}' = \{z'_w\}_{w\in\mathcal{W}'}$: if $\tau(x, z_{\varphi(w)}) = \tau(x', z'_w)$ for all $w \in \mathcal{W}'$, then $\mathcal{P}$ is valid.*

Intuitively, so long as complexity is "held fixed" across the two valuation tasks, the valuation paradigm is valid—even if the levels of valuations are distorted by systematic noise. Our specification of complexity τ yields sharp implications for when a given valuation paradigm satisfies this criterion. Below, we apply this criterion to the paradigms discussed in Example 2.

In what follows, suppose that $\succeq$ has an EU representation u and $\tau(\cdot,\cdot)$ has a CDF-complexity representation (u, H).

CE Risk Attitudes. This paradigm does not satisfy the criterion. Complexity is not held fixed across the valuation tasks: under CDF-complexity, valuing a lottery $x = (\overline{m}, p)$ against money is difficult for interior p whereas valuing its expected value $x' = (\overline{m} \cdot p, 1)$ is trivial, and so compression effects in the valuation of x can therefore bias this test of risk neutrality. Intuitively, as this test fundamentally relies on the levels of the valuation of x, it is not robust to systematic complexity-driven noise.

Paired CRP Test. It can be shown that this paradigm satisfies the criterion, and is therefore valid under CDF complexity. The intuition mirrors that of the invariance results in Proposition 1: under CDF complexity, the difficulty of valuing $x = (\overline{m}, p)$ against $z_w = (w, 1)$ is invariant to compounding x and prices by a common factor r, and so complexity is held fixed across $(x, \mathcal{Z})$ and $(x', \mathcal{Z}')$. As such, the paired valuation test of common ratio preferences proposed in McGranaghan et al. (2024b) is not only robust to the specifications of unbiased noise they consider, but also to our specification of complexity-driven noise.[12] Importantly, while the *comparisons* of paired valuations are valid under our model, it predicts that the *levels* of the valuations themselves—in particular, the inverse-S pattern McGranaghan et al. (2024b) document in both certainty equivalents and compounded certainty equivalents—may be confounded by complexity-driven noise.

Building on the tests of common ratio preferences in McGranaghan et al. (2024b), McGranaghan et al. (2024a) propose paired-valuation tests for common-consequence and mixture preferences, and document a range of aggregate violations of EU. In Appendix A.5, we show that these tests are also valid under CDF complexity.

Equalizing Reductions Tests. The equalizing reductions test of rank-dependence proposed in Bernheim and Sprenger (2020) satisfies the criterion, and so is valid in our framework. The intuition follows that of Proposition 2: under CDF complexity, the difficulty of valuing $x =$

[12]McGranaghan et al. (2024b) also show that their test of CRP is robust to systematic noise in valuations, under the assumption that the systematic component of noise is the same across paired valuation tasks. Our specification of noise provides foundations for this assumption.

$(a, b, c; p, q, 1-p-q)$ to $z_w = (a, b+m, c-w; p, q, 1-p-q)$ is invariant to the common payoff a; as such, complexity is held fixed across $(x, \mathcal{Z})$ and $(x', \mathcal{Z}')$.

On the other hand, the equalizing reductions test of probability weighting proposed in the same paper does not satisfy the criterion. Here, the difficulty of valuing $x = (a, b, c; p, q, 1-p-q)$ to $z_w = (a, b+m, c-w; p, q, 1-p-q)$ will in general vary with p and q under CDF-complexity; as discussed in Section 4.3, the resulting complexity-driven noise can lead to apparent non-linear probability weighting even under EU preferences.

Uncertainty Equivalents Test. The uncertainty equivalents paradigm satisfies the criterion, and so is a valid test of Independence under our framework. Intuitively, this is because the difficulty of valuing $x = (m_L, m_H; p, 1-p)$ against $z_w = (m_H; w)$ scales proportionally with p under CDF complexity. As a result, systematic noise induced by tradeoff-driven noise preserves the linear benchmark implied by EU, as shown in Proposition 3.

Andreoni and Sprenger (2011) provide a methodological argument for using uncertainty equivalents rather than certainty equivalents to study non-linear probability weighting: whereas inverse-S patterns in certainty equivalents may be hard to interpret, since under EU they can be confounded by exotic utility curvature, uncertainty equivalents provide a test of Independence that is robust to utility curvature. Our framework provides a complementary justification: complexity-driven noise can confound tests of inverse-S weighting under certainty equivalents, but it does not confound tests of Independence using uncertainty equivalents. In this sense, the uncertainty equivalents test is robust to both utility curvature and our specification of complexity-driven noise.

Using this valuation paradigm, Andreoni and Sprenger (2011) document violations of linearity near-certainty: in particular, for p' close to 1, they find that $\mathbb{E}[\varphi(V_{\mathcal{Z}}(x'))] > V_{\mathcal{Z}}(x)$. Because the uncertainty equivalents paradigm is valid in our framework, this pattern cannot be attributed to complexity-driven noise under EU, and instead points to a non-EU "certainty premium," consistent with the conclusions of Andreoni and Sprenger (2011).

6 Discussion

Motivated by evidence of systematic noise in valuations, this paper uses a formal model of complexity-driven noise to make progress on two methodological questions: how existing valuation data should be interpreted, and how valuation-based tests of preferences can be designed to mitigate the confound of complexity. While the focus of this paper is on the measurement of risk preferences, our modeling framework can be applied to other choice domains, such as the valuation of intertemporal prospects or multi-attribute goods, using the measures of tradeoff complexity developed for these domains in Shubatt and Yang (2024). As these measures quantify when complexity is held fixed across valuation tasks, our approach to complexity-robust valuation design can be similarly applied to motivate tests of preferences in these domains.

References

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, "Eliciting Risk and Time Preferences," *Econometrica*, May 2008, *76* (3), 583–618.

_ , _ , **Morten Igel Lau, and E. Elisabet Rutström**, "Elicitation using multiple price list formats," *Experimental Economics*, December 2006, *9* (4), 383–405.

Andreoni, James and Charles Sprenger, "Uncertainty Equivalents: Testing the Limits of the Independence Axiom," Technical Report w17342, National Bureau of Economic Research, Cambridge, MA August 2011.

Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, August 2018, *126* (4), 1472–1503.

Ballinger, T. Parker and Nathaniel T. Wilcox, "Decisions, Error and Heterogeneity," *The Economic Journal*, 1997, *107* (443), 1090–1105. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.1997.tb00009.x.

Beauchamp, Jonathan P., Daniel J. Benjamin, David I. Laibson, and Christopher F. Chabris, "Measuring and controlling for the compromise effect when estimating risk preference parameters," *Experimental Economics*, December 2020, *23* (4), 1069–1099.

Bell, David E., "Disappointment in Decision Making Under Uncertainty," *Operations Research*, February 1985, *33* (1), 1–27. Publisher: INFORMS: Institute for Operations Research.

Bernheim, B. Douglas and Charles Sprenger, "On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting," *Econometrica*, 2020, *88* (4), 1363–1409.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice Under Risk," *The Quarterly Journal of Economics*, August 2012, *127* (3), 1243–1285.

Bouchouicha, Ranoua, Ryan Oprea, Ferdinand M. Vieider, and Jilong Wu, "Is Prospect Theory Really a Theory of Choice?," 2024.

Bruner, David M., "Multiple switching behaviour in multiple price lists," *Applied Economics Letters*, April 2011, *18* (5), 417–420.

Cason, Timothy N. and Charles R. Plott, "Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing," *Journal of Political Economy*, December 2014, *122* (6), 1235–1270.

Diecidue, Enrico, Ulrich Schmidt, and Peter P. Wakker, "The Utility of Gambling Reconsidered," *Journal of Risk and Uncertainty*, December 2004, *29* (3), 241–259.

Echenique, Federico and Gerelt Tserenjigmid, "Robust Testing Of the Allais Paradox By Paired Choices vs. Paired Valuations," April 2026. arXiv:2604.06050 [econ].

Enke, Benjamin and Cassidy Shubatt, "Quantifying Lottery Choice Complexity," September 2023.

_ and Thomas Graeber, "Cognitive Uncertainty," *The Quarterly Journal of Economics*, September 2023, *138* (4), 2021–2067.

_ , _ , Ryan Oprea, and Jeffrey Yang, "Behavioral Attenuation," September 2024.

Feldman, Paul J. and Paul J. Ferraro, "Certain Reversals," 2024.

Freeman, David J. and Guy Mayraz, "Why choice lists increase risk taking," *Experimental Economics*, March 2019, *22* (1), 131–154.

_ , Yoram Halevy, and Terri Kneeland, "Eliciting risk preferences using choice lists," *Quantitative Economics*, 2019, *10* (1), 217–237.

Frydman, Cary and Lawrence J Jin, "On the source and instability of probability weighting," 2025.

Gul, Faruk, "A Theory of Disappointment Aversion," *Econometrica*, 1991, *59* (3), 667–686.

Hardisty, David J., Katherine F. Thompson, David H. Krantz, and Elke U. Weber, "How to measure time preferences: An experimental comparison of three methods," *Judgment and Decision Making*, May 2013, *8* (3), 236–249.

Harless, David W. and Colin F. Camerer, "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 1994, *62* (6), 1251–1289.

Harrison, G. W., M. I. Lau, E. E. Rutström, and M. B. Sullivan, "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues," *Field Experiments in Economics*, 2005.

Holden, Stein T., Sarah Tione, Mesfin Tilahun, and Samson Katengeza, "Elicitation bias in Multiple Price Lists: A field experiment," *Journal of Behavioral and Experimental Economics*, 2025, p. 102479.

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

Imai, Taisuke, Salvatore Nunnari, Jilong Wu, and Ferdinand M. Vieider, "Meta-Analysis of Prospect Theory Parameters," 2025.

Jack, B. Kelsey, Kathryn McDermott, and Anja Sautmann, "Multiple price lists for willingness to pay elicitation," *Journal of Development Economics*, November 2022, *159*, 102977.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica (pre-1986)*, March 1979, *47* (2), 263–291. Num Pages: 29 Place: Evanston, United Kingdom.

Karni, Edi and Zvi Safra, ""Preference Reversal" and the Observability of Preferences by Experimental Methods," *Econometrica*, May 1987, *55* (3), 675.

Kontek, Krzysztof, "Boundary effects in the Marschak-Machina triangle," *Judgment and Decision Making*, November 2018, *13* (6), 587–606.

— **and Michal Lewandowski**, "Range-Dependent Utility," *Management Science*, June 2018, *64* (6), 2812–2832.

Loomes, Graham and Robert Sugden, "Disappointment and Dynamic Consistency in Choice under Uncertainty," *The Review of Economic Studies*, April 1986, *53* (2), 271.

— **and** — , "Testing Different Stochastic Specificationsof Risky Choice," *Economica*, 1998, *65* (260), 581–598. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0335.00147.

McGranaghan, Christina, Kirby Nielsen, Ted O'Donoghue, Jason Somerville, and Charles D Sprenger, "Connecting Common Ratio and Common Consequence Preferences," September 2024.

— , — , — , — , **and Charles D. Sprenger**, "Distinguishing Common Ratio Preferences from Common Ratio Effects Using Paired Valuation Tasks," *American Economic Review*, February 2024, *114* (2), 307–347.

Natenzon, Paulo, "Random Choice and Learning," *Journal of Political Economy*, February 2019.

Neilson, William S., "Some mixed results on boundary effects," *Economics Letters*, July 1992, *39* (3), 275–278.

O'Donoghue, Ted and Jason Somerville, "Investigating Risk Preferences Using Experiments," in Alex Rees-Jones, ed., *Handbook of Experimental Methods in the Social Sciences*, Vol. 1, Edward Elgar Publishing, 2026. Status: in press.

Oprea, Ryan, "Decisions under risk are decisions under complexity," *American Economic Review*, 2024, *114* (12), 3789–3811.

Schmidt, Ulrich, "A Measurement of the Certainty Effect," *Journal of Mathematical Psychology*, March 1998, *42* (1), 32–47.

Seidl, Christian, "Preference reversal," *Journal of Economic Surveys*, 2002, *16* (5), 621–655.

Shubatt, Cassidy and Jeffrey Yang, "Tradeoffs and Comparison Complexity," January 2024. arXiv:2401.17578 [econ].

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, December 2015, *123* (6), 1456–1499.

Tversky, Amos and Daniel Kahneman, "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, October 1992, *5* (4), 297–323.

_ **and Richard H Thaler**, "Anomalies: Preference Reversals," *Journal of Economic Perspectives*, May 1990, *4* (2), 201–211.

_ **, Paul Slovic, and Daniel Kahneman**, "The Causes of Preference Reversal," *The American Economic Review*, 1990, *80* (1), 204–217.

APPENDIX

A Additional Theoretical Results

A.1 Microfoundation of Signal Structure

We show that our model of complexity-driven noise in valuations can be microfounded as a limit case of the choice model of imprecise comparisons developed in Shubatt and Yang (2024) when applied to multiple price list valuations.

Valuations in a multiple price list. In this model, the DM has continuous, iid priors with distribution Q over the utility of each lottery in $x \in X$, denoted u_x, where u_x is assumed to represent a proper preference $\succeq$. In the simplest form of the model, the DM is faced with a finite menu of options A, receives signals for s_{xy} on each ordinal value comparison between options $x, y \in A$, integrates these signals using Bayes' rule, and then chooses the option with highest posterior expected value (randomizing in the case of ties). We consider signals of the following form:

$$s_{xy} = \text{sign}(u_x - u_y)\epsilon_{xy},$$
$$\epsilon_{xy} \sim \text{Bernoulli}(\tau(x, y))$$

where $\tau(\cdot, \cdot)$ is a comparability function, and the ϵ_{xy} are independent. That is, the signal is fully revealing of the ordinal value comparison between x and y (if there is a strict preference) with probability $\tau(x, y)$, and is uninformative otherwise.[13] In this model, a DM is represented by (u, τ, Q).

Shubatt and Yang (2024) apply the model to multiple price list valuations as follows. For a finite set of numeraire amounts $W \subset \mathbb{R}$, a *price list* is a family of lotteries indexed by W, $Z = \{z_w\}_{w \in W}$, where $z_w <_{FOSD} z_{w'}$ if and only if $w < w'$. We will enumerate the elements of W in increasing order by $\{w_1, w_2, ..., w_n\}$. An *MPL valuation task* (x, Z) is a pairing of a lottery x and price list Z, where there exist $z, z' \in Z$ such that $u_z \leq u_x \leq u_{z'}$.

When faced with an MPL valuation task (x, Z), the DM receives signals s_{xy} for each pairwise comparison in $\{x\} \cup Z$, integrates these signals using Bayes' rule, and then chooses the option with the highest posterior expected value from each binary menu $(\{x, z_w\})_{w \in W}$ (again randomizing in the case of ties). Under any signal realization, these decisions induce a switching point $R \in \{1, ..., n-1\}$, where she chooses $x \in \{x, z_{w_k}\}$ for $k \leq R$ and $z_{w_k} \in \{x, z_{w_k}\}$ for $k > R$. The

[13] Shubatt and Yang (2024) consider ordinal signals with a continuous error structure. It can be shown that all formal and numerical results therein have direct analogs under the error structure we consider here.

model makes predictions about the distribution of switching points, denoted $R_Z(x)$.[14] Associate each switching point R with the valuation $\frac{1}{2}(w_R + w_{R+1})$, i.e. the midpoint of the adjacent prices; let $V_Z(x)$ denote the distribution over these valuations induced by $R_Z(x)$.[15]

Limit result. Our model of valuations in a price structure $(x, \mathcal{Z})$ can be derived as the limit case of such valuations from a price list (x, Z) as the list becomes increasingly fine-grained.

For simplicity, we derive this limit result in the case of a bounded valuation task $(x, \mathcal{Z})$. Define as before $\underline{w}_{\mathcal{Z}}(x) = \inf\{w \in \mathcal{W} : z_w \not<_{FOSD} x\}$ and $\overline{w}_{\mathcal{Z}}(x) = \sup\{w \in \mathcal{W} : z_w \not>_{FOSD} x\}$ the range of undominated prices. For $n > 2$, let $Z^n = \{z_{w_1}, ..., z_{w_n}\}$ denote the *adapted* price list, which satisfies $w_1 = \underline{w}_{\mathcal{Z}}(x)$, $w_n = \underline{w}_{\mathcal{Z}}(x)$, and $w_{k+1} - w_k$ constant in k. That is, an adapted price list Z^n consists of evenly-spaced prices in $[\underline{w}_{\mathcal{Z}}(x), \overline{w}_{\mathcal{Z}}(x)]$.

Proposition 5. *Let $V_{\mathcal{Z}}(x)$ denote the valuation associated with proper preference $\succeq$ and comparability function τ, where $t \mapsto \ln(1 - \tau(x, z_t))^2$ is integrable on $[\underline{w}, \overline{w}]$. Let $V_{Z^n}(x)$ denote the multiple price list valuation of a DM with (u, τ^n, Q), where $\tau^n(\cdot,\cdot) = 1 - (1 - \tau(\cdot,\cdot))^{1/(n-1)}$ and u represents $\succeq$. If Q is uniform, then $V_{Z^n}(x) \xrightarrow{D} V_{\mathcal{Z}}(x)$.*

In words, as the price list becomes increasingly fine-grained, and as the informativeness of any individual comparison (x, z_w) is scaled down commensurately, multiple price list valuations converge to the error structure we consider in the main text.

A.2 Monotonicity in Valuations

Recall that a comparability function τ is monotone if $x \succeq y$, then $x' >_{FOSD} x$ implies $\tau(x', y) > \tau(x, y)$ and $y >_{FOSD} y'$ implies $\tau(x, y) < \tau(x, y')$. The comparability functions we consider in this paper, such as CDF-complexity, satisfy monotonicity. The result below states that monotonicity in τ (as well as underlying preferences) implies a form of monotonicity in valuations.

Proposition 6. *Suppose preferences $\succeq$ are proper and the comparability function τ is monotone. Then, for valuation tasks $(x, \mathcal{Z})$ and $(x', \mathcal{Z})$ where $x' >_{FOSD} x$ and $\overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x') = \overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)$, we have $V_{\mathcal{Z}}(x') >_{FOSD} V_{\mathcal{Z}}(x)$.*

Notice that in addition to monotonicity of τ, we also require $\overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x') = \overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)$ to ensure that valuations respect monotonicity. This is because we normalize the rate functions $\lambda^x_{\mathcal{Z}}(w)$, $\lambda^{x'}_{\mathcal{Z}}(w)$ by the range of undominated prices; $\overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x') = \overline{w}_{\mathcal{Z}}(x) -$

[14]Signals may lead to ties, where there exists some $R \in \{1, ..., n\}$ such that the DM chooses $x \in \{x, z_{w_k}\}$ for all $k < R$ and $z_{w_k} \in \{x, z_{w_k}\}$ for all $k > R$, and randomizes between $\{x, z_{w_R}\}$. In this case, we define the switching point as this R.

[15]If the switching point R arises from a tie, we instead define the valuation to be w_R.

$\underline{w}_{\mathcal{Z}}(x)$ ensures that $\lambda^{x}_{\mathcal{Z}}(w)$, $\lambda^{x'}_{\mathcal{Z}}(w)$ employ the same normalization, and so monotonicity in τ guarantees monotonicity in the rate functions.

A.3 Extension to Rank-Dependent Preferences

A *probability weighting function* is a function $\pi : [0,1] \to [0,1]$ that is weakly increasing, continuous, and satisfies $\pi(0) = 0$ and $\pi(1) = 1$. In rank-dependent utility, the utility of a lottery x is given by

$$RDU(x) = \int_0^1 u(F_x^{-1}(q))\, d\hat{\pi}(q)$$

where $u : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function and $\hat{\pi}(q) = 1 - \pi(1-q)$, for π a probability weighting function. Say that $\succeq$ has a *rank-dependent utility* (RDU) representation (u, π) if it is represented by the utility function above. We extend the CDF-complexity measure to such preferences.

Definition 3. *A comparability function τ has a generalized CDF-complexity representation (u, π, H) if there exists strictly increasing u and probability weighting function π such that for all $x \neq y$,*

$$\tau(x, y) = H\left(\frac{|RDU(x) - RDU(y)|}{d_{CDF}(x, y)}\right)$$

for H strictly increasing with $H(0) = 0, H(1) = 1$, where $RDU(x) = \int_0^1 u(F_x^{-1}(q))\, d\hat{\pi}(q)$ for $\hat{\pi}(q) = 1 - \pi(1-q)$ and

$$d_{CDF}(x, y) = \int_0^1 |u(F_x^{-1}(q)) - u(F_y^{-1}(q))|\, d\hat{\pi}(q).$$

This definition is identical to the standard CDF-complexity measure in Definition 1, except the quantile functions in the numerator and denominator are integrated over using a non-uniform measure determined by the probability weighting function π; when π is linear, the representation reduces to Definition 1. Notice that this generalization shares several core properties with CDF-complexity, chiefly dominance and monotonicity, but relaxes linearity.

A.4 Validity for Sign Tests

Consider an alternative definition of validity that corresponds to a sign test.

Definition 4. *Say that a valuation paradigm $((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$ is sign-valid if whenever $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$, $Med(V_{\mathcal{Z}}(x)) = Med(\varphi(V_{\mathcal{Z}'}(x')))$.*

Proposition 4 gives sufficient conditions for a valuation paradigm to be valid if the analyst is interested in testing a comparison of means. The exact same conditions guarantee sign-validity.

Proposition 7. *Consider a valuation paradigm* $\mathcal{P} = ((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$, *where* $\varphi(\underline{w}_{\mathcal{Z}'}(x')) = \underline{w}_{\mathcal{Z}}(x)$ *and* $\varphi(\overline{w}_{\mathcal{Z}'}(x')) = \overline{w}_{\mathcal{Z}}(x)$. *Denoting* $\mathcal{Z} = \{z_w\}_{w \in \mathcal{W}}$ *and* $\mathcal{Z}' = \{z'_w\}_{w \in \mathcal{W}'}$: *if* $\tau(x, z_{\varphi(w)}) = \tau(x', z'_w)$ *for all* $w \in \mathcal{W}'$, *then* $\mathcal{P}$ *is sign-valid.*

A.5 Additional Paired Valuation Tests

McGranaghan et al. (2024a) develop a framework for testing common-ratio preferences, common-consequence preferences, and mixture preferences using paired valuation tests. Fixing $m > 0$, $p, r \in (0, 1)$, consider the following valuation paradigms:

- *Paired CRP Test – Alternative*: $x = (m; 1)$, $\mathcal{Z} = \{(w; p)\}_{w \in [m, \overline{w}]}$, $x' = (m; r)$, $\mathcal{Z}' = \{(w; rp)\}_{w \in [m, \overline{w}]}$, $\varphi(v) = v$. Here, $v_{\mathcal{Z}}(x) > v_{\mathcal{Z}'}(x')$ is indicative of a common ratio preference.
- *Paired CCP Test*: $x = (m; 1)$, $\mathcal{Z} = \{(w, m; rp, 1-r)\}_{w \in [m, \overline{w}]}$, $x' = (m; r)$, $\mathcal{Z}' = \{(w; rp)\}_{w \in [m, \overline{w}]}$, $\varphi(v) = v$. Here, $v_{\mathcal{Z}}(x) > v_{\mathcal{Z}'}(x')$ is indicative of a common consequence preference.
- *Paired MXP Test*: $x = (m; 1)$, $\mathcal{Z} = \{(w; p)\}_{w \in [m, \overline{w}]}$, $x' = (m; 1)$, $\mathcal{Z}' = \{(w, m; pr, 1 - r)\}_{w \in [m, \overline{w}]}$, $\varphi(v) = v$. Here, $v_{\mathcal{Z}}(x) > v_{\mathcal{Z}'}(x')$ is indicative of a preference for mixtures.

It is easy to see that if $\succeq$ is EU and τ has a CDF-complexity representation, all three of these valuation paradigms satisfy the criterion in Proposition 4, and are therefore valid. This is because in all three paradigms, the lotteries in $(x', \mathcal{Z}')$ are obtained by mixing the lotteries in $(x, \mathcal{Z}')$ with a common payoff, or by removing a common consequence from the lotteries in $(x, \mathcal{Z}')$; under CDF complexity, this preserves the difficulty of comparing the lottery to prices.

A.6 Context-Dependent Preferences

Our experiment documents apparent inconsistencies in lottery preference across valuation formats. Of course, we are not the first to document and seek to rationalize unstable revealed preferences across experimental paradigms. Past work has proposed theories of context-dependent preferences, wherein changes in context can explain inconsistencies across experimental paradigms. Below, we discuss several of the most prominent theories of context-dependent preferences and explain why they are unable to generate the patterns observed in our experiment.

Salience Theory. In salience theory of choice under risk (Bordalo et al., 2012) the value of a good is dependent on the menu in which the good is available. For lotteries in particular, the menu of choices determines which states of lotteries are most salient, and salient states are overweighted when assessing value.

Salience offers an explanation for the revealed preference inconsistencies observed in classic preference reversal experiments (e.g. Tversky and Thaler, 1990; Seidl, 2002) wherein given a safer lottery ℓ_π which has a high probability of a low payoff, and a riskier lottery $\ell_\$$ which has a low probability of a high payoff, subjects favor ℓ_π in direct choice between the two, yet tend to assign higher valuations to $\ell_\$$ when valuing the lotteries in terms of certainty equivalents. Bordalo et al. (2012) explain this finding as follows. In direct choice, where the menu is the pair of both lotteries, the state in which $\ell_\$$ loses and ℓ_π wins is highly salient, producing a preference for ℓ_π. When lotteries are instead priced in isolation, the menu is the lottery and the alternative of having 0 for sure. Bordalo et al. (2012) show that this alternative reference can generate a *higher* value for $\ell_\$$ and a *lower* value for ℓ_π, compared to when in a menu together.

This kind of menu instability cannot explain inconsistencies across valuation tasks, if we are committed to the paper's assumption that the appropriate alternative when considering a lottery in isolation is having zero for sure. This alternative would be the same across each of the different valuation formats we consider, and so cannot drive inconsistent preferences across these formats. At the same time, as Shubatt and Yang (2024) show, our model of tradeoff-driven noise in valuations can explain classic preference reversals, under the simple assumption that tradeoff difficulty $\tau(\cdot,\cdot)$ produces unbiased noise in direct comparisons.

Expectations-Based Reference Dependence. Reference dependence models propose a different form of context dependence, where value is assessed relative to some reference point which may change across contexts. In these models, losses relative to the reference point are weighed more heavily than gains. Sprenger (2015) and Feldman and Ferraro (2024) show that models of expectations based reference dependence (EBRD) generate patterns of valuations across certainty equivalents ($\ell = (\overline{m}; p)$ valued against $\mathcal{Z}_{CE} = \{(w; 1)_{w\in\mathbb{R}}\}$) and probability equivalents ($c = (m; 1)$ valued against $\mathcal{Z}_{PE} = \{(\overline{m}; w)\}_{w\in[0,1]}$).

In the formulation of expectations-based reference dependence of Sprenger (2015), the reference point used in a valuation task is the option being valued: so when a lottery is valued against certain payments, the reference point is the lottery; and when a certain payment is valued against lotteries, the reference point is the certain payment. Sprenger (2015) shows formally that under this assumption, EBRD predicts that valuations are more risk-averse in probability equivalents than in certainty equivalents. This model similarly makes the prediction that valuations should be more risk averse in our probability equivalent tasks (P-a and P-b) relative to our money equivalent tasks (M-a and M-b). This rules out several patterns we observe in our data; for instance, cases where P-a valuations are risk-seeking while M-a valuations are risk-averse. Feldman and Ferraro (2024) augment the model of Sprenger (2015) by allowing value function curvature to depend on the valuation context, and in particular assume greater value function convexity in money equivalents, where monetary payoffs vary across prices, relative to probability equivalents, where the monetary payoff is fixed across prices. As with Sprenger (2015), this model cannot easily account for risk seeking P-a valuations for high values of p.

A.7 Restrictions on Valuation Reversals

Consider the experimental environment in Section 2, where $\ell_R = (30; 0.2 \cdot p)$, $\ell_S = (30 \cdot p; 0.2)$ are valued against four distinct price structures: $\mathcal{Z}_{M\text{-}a} = \{(w; 0.2)\}_{w \in \mathbb{R}}$, $\mathcal{Z}_{M\text{-}b} = \{(w; 0.5)\}_{w \in \mathbb{R}}$, $\mathcal{Z}_{P\text{-}a} = \{(30; w)\}_{w \in [0,1]}$, and $\mathcal{Z}_{P\text{-}b} = \{(40; w)\}_{w \in [0,1]}$. For $F \in \{M\text{-}a, M\text{-}b, P\text{-}a, P\text{-}b\}$, let $\succeq_F$ denote the modal preference over lotteries under the price structure $\mathcal{Z}_F$, where $\ell \succeq_F \ell'$ whenever $\mathbb{E}[V_{\mathcal{Z}_F}(\ell)] \geq \mathbb{E}[V_{\mathcal{Z}_F}(\ell')]$, and define the strict modal preference $\succ_F$ analogously.

In our experiment, we documented inconsistencies in modal preferences across our four price strictures, summarized in Table 3. In Figure 4, we showed that our parameterized valuation model, where preferences are given by $u(x) = x^\alpha$ and $\tau(\cdot, \cdot)$ has a CDF-complexity representation (u, H) with $H(r) = 1 - (1 - r)^\gamma$, can rationalize each of these inconsistencies. Here we provide formal and numerical results to show that the parameterized model is restrictive, in the sense that for each documented inconsistency our model permits, the *opposite* inconsistency is ruled out by the model.

$p \in \{0.1, 0.2\}$	$p \in \{0.8, 0.9\}$
$\ell_R \succ_{M\text{-}a} \ell_S$, $\ell_R \prec_{P\text{-}a} \ell_S$	$\ell_R \prec_{M\text{-}a} \ell_S$, $\ell_R \succ_{P\text{-}a} \ell_S$
$\ell_R \succ_{M\text{-}a} \ell_S$, $\ell_R \prec_{P\text{-}b} \ell_S$	$\ell_R \prec_{M\text{-}a} \ell_S$, $\ell_R \succ_{M\text{-}b} \ell_S$
$\ell_R \prec_{P\text{-}a} \ell_S$, $\ell_R \succ_{M\text{-}b} \ell_S$	$\ell_R \succ_{P\text{-}a} \ell_S$, $\ell_R \prec_{P\text{-}b} \ell_S$
$\ell_R \succ_{M\text{-}b} \ell_S$, $\ell_R \prec_{P\text{-}b} \ell_S$	$\ell_R \succ_{M\text{-}b} \ell_S$, $\ell_R \prec_{P\text{-}b} \ell_S$

Table 3: Modal preference inconsistencies over the safe lottery ℓ_s and the risky lottery ℓ_R documented in the experiment.

First, we show formally that the model rules out the opposite inconsistencies between M-a valuations and both P-a and P-b valuations.

Proposition 8. *Suppose $\succeq$ has an EU representation $u(x) = x^\alpha$ and $\tau(\cdot, \cdot)$ has a CDF-complexity representation (u, H) with $H(r) = 1 - (1 - r)^\gamma$. Then*

(a) For $p \in \{0.1, 0.2\}$, $\ell_R \preceq_{M\text{-}a} \ell_S$ implies $\ell_R \prec_{P\text{-}a} \ell_S$, and for $p \in \{0.8, 0.9\}$, $\ell_R \succeq_{M\text{-}a} \ell_S$ implies $\ell_R \succ_{P\text{-}a} \ell_S$.

(b) For $p \in \{0.1, 0.2\}$, $\ell_R \preceq_{M\text{-}a} \ell_S$ implies $\ell_R \prec_{P\text{-}b} \ell_S$

The following result gives a sufficient condition under which the model rules out the opposite inconsistencies between all other valuation formats.

Proposition 9. *Suppose $\succeq$ has an EU representation $u(x) = x^\alpha$ and $\tau(\cdot,\cdot)$ has a CDF-complexity representation (u, H) with $H(r) = 1-(1-r)^\gamma$. If for all $\gamma > 0$ and $p \in \{0.1, 0.2, 0.8, 0.9\}$, we have $\ell_R \succ_{M\text{-}b} \ell_S$ for all $\alpha \geq 1$ and $\ell_R \prec_{M\text{-}b} \ell_S$ for all $\alpha \leq 1$, then*

(a) For $p \in \{0.8, 0.9\}$, $\ell_R \succeq_{M\text{-}a} \ell_S$ implies $\ell_R \succ_{M\text{-}b} \ell_S$.

(b) For $p \in \{0.1, 0.2\}$, $\ell_R \succeq_{P\text{-}a} \ell_S$ implies $\ell_R \succ_{M\text{-}b} \ell_S$.

(c) For $p \in \{0.8, 0.9\}$, $\ell_R \preceq_{P\text{-}a} \ell_S$ implies $\ell_R \prec_{P\text{-}b} \ell_S$.

(d) For $p \in \{0.1, 0.2, 0.8, 0.9\}$, $\ell_R \preceq_{M\text{-}b} \ell_S$ implies $\ell_R \prec_{P\text{-}b} \ell_S$.

We provide numerical evidence for this condition. For $p \in \{0.1, 0.2, 0.8, 0.9\}$, we compute $\Delta_{M\text{-}b}(p) \equiv \mathbb{E}[V_{\mathcal{Z}_{M\text{-}b}}(\ell_R)] - \mathbb{E}[V_{\mathcal{Z}_{M\text{-}b}}(\ell_S)]$ and $\Delta_{P\text{-}b}(p) = \mathbb{E}[V_{\mathcal{Z}_{P\text{-}b}}(\ell_R)] - \mathbb{E}[V_{\mathcal{Z}_{P\text{-}b}}(\ell_S)]$ for the parameter grid $\alpha \in \{0.01, 0.02, ..., 0.99, 1, 1/0.99, ..., 1/0.02, 1/0.01\}$, $\gamma \in \{0.1, 0.2, ..., 19.9, 20\}$. We find that for all values of α greater than or equal to 1, the minimum value of $\Delta_{M\text{-}b}(p)$ is 0.0962, and for all values of α less than or equal to 1, the maximum value of $\Delta_{P\text{-}b}(p)$ is -0.0006.

B Proofs

It will be useful to observe the following results.

Lemma 1. *Consider a valuation paradigm $\mathcal{P} = ((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$, where $\varphi(\underline{w}_{\mathcal{Z}'}(x')) = \underline{w}_{\mathcal{Z}}(x)$ and $\varphi(\overline{w}_{\mathcal{Z}'}(x')) = \overline{w}_{\mathcal{Z}}(x)$. Denoting $\mathcal{Z} = \{z_w\}_{w\in\mathcal{W}}$ and $\mathcal{Z}' = \{z'_w\}_{w\in\mathcal{W}'}$: if $\tau(x, z_{\varphi(w)}) = \tau(x', z'_w)$ for all $w \in \mathcal{W}'$, then $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$ implies $V_{\mathcal{Z}}(x) =_D \varphi(V_{\mathcal{Z}'}(x'))$*

Proof. Since φ is affine, we can represent it as $\varphi(x) = \alpha x + \beta$. Since by hypothesis $\overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x) = \varphi(\overline{w}_{\mathcal{Z}'}(x')) - \varphi(\underline{w}_{\mathcal{Z}'}(x'))$, we have $\alpha = \frac{\overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)}{\overline{w}_{\mathcal{Z}'}(x') - \underline{w}_{\mathcal{Z}'}(x')}$.

Now suppose that $v_{\mathcal{Z}}(x) = \varphi(v_{\mathcal{Z}'}(x'))$. Since $V_{\mathcal{Z}}(x) = v_{\mathcal{Z}}(x) + (\overline{\delta}_{\mathcal{Z}}(x) - \underline{\delta}_{\mathcal{Z}}(x))/2$, we have $\varphi^{-1}(V_{\mathcal{Z}}(x)) = v_{\mathcal{Z}'}(x') + (\overline{\delta}_{\mathcal{Z}}(x)/\alpha + \underline{\delta}_{\mathcal{Z}}(x)/\alpha)/2$. Note that

$$\mathbb{P}(\overline{\delta}_{\mathcal{Z}}(x)/\alpha \leq t) = \begin{cases} 1 - \exp(-\int_0^{\alpha t} \lambda^x_{\mathcal{Z}}(v_{\mathcal{Z}}(x) + r)\, dr & \alpha t < \overline{w}_{\mathcal{Z}}(x) - v_{\mathcal{Z}}(x) \\ 1 & \text{otherwise} \end{cases}$$

$$= \begin{cases} 1 - \exp(-\int_0^{\alpha t} \lambda^x_{\mathcal{Z}}(\varphi(v_{\mathcal{Z}'}(x')) + r)\, dr & t < \overline{w}_{\mathcal{Z}'}(x') - v_{\mathcal{Z}'}(x') \\ 1 & \text{otherwise} \end{cases}.$$

By a change of variables $\tilde{r} = r/\alpha$, we have

$$\begin{aligned}\int_0^{\alpha t} \lambda_{\mathcal{Z}}^x(\varphi(v_{\mathcal{Z}'}(x')) + r)\,dr &= \int_0^t \alpha\lambda_{\mathcal{Z}}^x(\varphi(v_{\mathcal{Z}'}(x')) + \alpha r)\,dr \\ &= \int_0^t \alpha\lambda_{\mathcal{Z}}^x(\varphi(v_{\mathcal{Z}'}(x') + r))\,dr \\ &= \int_0^t \lambda_{\mathcal{Z}'}^{x'}(v_{\mathcal{Z}'}(x') + r)\,dr\end{aligned}$$

where the final equality follows from the fact that $\alpha = \frac{\overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)}{\overline{w}_{\mathcal{Z}'}(x') - \underline{w}_{\mathcal{Z}'}(x')}$ and $\tau(x, z_{\varphi(v_{\mathcal{Z}}(x)+r)}) = \tau(x', z_{v_{\mathcal{Z}'}(x')+r})$. This implies that $\overline{\delta}_{\mathcal{Z}}(x)/\alpha =_D \overline{\delta}_{\mathcal{Z}'}(x')$. An analogous argument shows that $\underline{\delta}_{\mathcal{Z}}(x)/\alpha =_D \underline{\delta}_{\mathcal{Z}'}(x')$; independence of $\overline{\delta}_{\mathcal{Z}}(x), \underline{\delta}_{\mathcal{Z}}(x)$ therefore implies that $\varphi^{-1}(V_{\mathcal{Z}}(x)) =_D V_{\mathcal{Z}'}(x')$ as desired. □

For the following result, let $V_{\mathcal{Z}}(x)$ and $v_{\mathcal{Z}}(x)$ denote the observed and true valuations of a DM described by $(\succeq, \tau(\cdot,\cdot))$, and let $V'_{\mathcal{Z}}(x)$ and $v'_{\mathcal{Z}}(x)$ denote the observed and true valuations of a DM described by $(\succeq', \tau'(\cdot,\cdot))$

Lemma 2. *For a price structure $\mathcal{Z} = \{z_w\}_{w\in\mathcal{W}}$, consider valuation tasks $(x, \mathcal{Z})$ and $(x', \mathcal{Z})$ satisfying $\overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x) = \overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x')$, $\overline{w}_{\mathcal{Z}}(x') > \overline{w}_{\mathcal{Z}}(x)$, and $\underline{w}_{\mathcal{Z}}(x') > \underline{w}_{\mathcal{Z}}(x)$. If $v'_{\mathcal{Z}}(x') > v_{\mathcal{Z}}(x)$, $\tau'(x', z_w) < \tau(x, z_w)$ for all $w \geq V'_{\mathcal{Z}}(x')$, and $\tau'(x', z_w) > \tau(x, z_w)$ for all $w < v_{\mathcal{Z}}(x)$, then $V'_{\mathcal{Z}}(x') >_{FOSD} V_{\mathcal{Z}}(x)$.*

Proof. For $y \in \{x, x'\}$, define $\overline{V}_{\mathcal{Z}}(y) = v_{\mathcal{Z}}(y) + \overline{\delta}_{\mathcal{Z}}(y)$ and $\underline{V}_{\mathcal{Z}}(y) = v_{\mathcal{Z}}(y) - \underline{\delta}_{\mathcal{Z}}(y)$, and define $\overline{V}'_{\mathcal{Z}}(y)$ and $\underline{V}'_{\mathcal{Z}}(y)$ analogously; we have $V_{\mathcal{Z}}(y) = (\overline{V}_{\mathcal{Z}}(y) + \underline{V}_{\mathcal{Z}}(y))/2$ and $V'_{\mathcal{Z}}(y) = (\overline{V}'_{\mathcal{Z}}(y) + \underline{V}'_{\mathcal{Z}}(y))/2$. Since $\underline{V}_{\mathcal{Z}}(y)$ and $\overline{V}_{\mathcal{Z}}(y)$ are independent, and $\underline{V}'_{\mathcal{Z}}(y)$ and $\overline{V}'_{\mathcal{Z}}(y)$ are independent, it suffices to show that $\overline{V}'_{\mathcal{Z}}(x') >_{FOSD} \overline{V}_{\mathcal{Z}}(x)$ and $\underline{V}'_{\mathcal{Z}}(x') >_{FOSD} \underline{V}_{\mathcal{Z}}(x)$.

For $y \in x, x'$, let $\underline{G}_y$ and $\overline{G}_y$ denote the distributions of $\underline{V}_{\mathcal{Z}}(y)$ and $\overline{V}_{\mathcal{Z}}(y)$, respectively, and define $\underline{G}'_y$ and $\overline{G}'_y$ analogously. We have

$$\overline{G}_y(t) = \begin{cases} 0 & t < v_{\mathcal{Z}}(y) \\ 1 - \exp\left(-\int_{v_{\mathcal{Z}}(y)}^t \lambda_{\mathcal{Z}}^y(r)\,dr\right) & t \in [v_{\mathcal{Z}}(y), \overline{w}_{\mathcal{Z}}(y)) \\ 1 & \text{otherwise} \end{cases}$$

$$\underline{G}_y(t) = \begin{cases} 0 & t > v_{\mathcal{Z}}(y) \\ \exp\left(-\int_t^{v_{\mathcal{Z}}(y)} \lambda_{\mathcal{Z}}^y(r)\,dr\right) & t \in (\underline{w}_{\mathcal{Z}}(y), v_{\mathcal{Z}}(y)] \\ 1 & \text{otherwise} \end{cases}$$

where $\lambda^y_{\mathcal{Z}}(t) = -\frac{1}{\overline{w}_{\mathcal{Z}}(y) - \underline{w}_{\mathcal{Z}}(y)} \cdot \ln(1-\tau(x,z_t))$. $\underline{G}'_y$ and $\overline{G}'_y$ take the analogous form with respect to the rate function $\lambda'^y_{\mathcal{Z}}(t) = -\frac{1}{\overline{w}_{\mathcal{Z}}(y) - \underline{w}_{\mathcal{Z}}(y)} \cdot \ln(1-\tau'(x,z_t))$.

Since $\overline{w}_{\mathcal{Z}}(x) \leq \overline{w}_{\mathcal{Z}}(x')$ and $v'_{\mathcal{Z}}(x') > v_{\mathcal{Z}}(x)$, we have $\overline{G}'_{x'}(w) \leq \overline{G}_x(w)$ for $w < v'_{\mathcal{Z}}(x')$ and $w > \overline{w}_{\mathcal{Z}}(x)$. Now fix $w \in [v'_{\mathcal{Z}}(x'), \overline{w}_{\mathcal{Z}}(x)]$; by assumption $\tau'(x', z_w) < \tau(x, z_w)$. Since by assumption $\overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x') = \overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)$, we have $\lambda'^{x'}_{\mathcal{Z}}(w) < \lambda^x_{\mathcal{Z}}(w)$, which in turn implies that $\int_{v_{\mathcal{Z}}(x')}^{w} \lambda'^{x'}_{\mathcal{Z}}(r)\,dr < \int_{v_{\mathcal{Z}}(x)}^{w} \lambda^x_{\mathcal{Z}}(r)\,dr$. Therefore, $\overline{G}'_{x'}(w) < \overline{G}_x(w)$ for all $w \in [v'_{\mathcal{Z}}(x'), \overline{w}_{\mathcal{Z}}(x)]$, and so $\overline{V}'_{\mathcal{Z}}(x') >_{FOSD} \overline{V}_{\mathcal{Z}}(x)$.

Since $\underline{w}_{\mathcal{Z}}(x') \leq \underline{w}_{\mathcal{Z}}(x)$ and $v'_{\mathcal{Z}}(x') > v_{\mathcal{Z}}(x)$, we have $\underline{G}_{x'}(w) \leq \underline{G}_x(w)$ for $w > v_{\mathcal{Z}}(x)$ and $w < \underline{w}_{\mathcal{Z}}(x')$. Now fix $w \in [\underline{w}_{\mathcal{Z}}(x'), v_{\mathcal{Z}}(x)]$; by assumption $\tau'(x', z_w) > \tau(x, z_w)$. Since by assumption $\overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x') = \overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)$, we have $\lambda'^{x'}_{\mathcal{Z}}(w) > \lambda^x_{\mathcal{Z}}(w)$, which in turn implies that $\int_{w}^{v_{\mathcal{Z}}(x')} \lambda'^{x'}_{\mathcal{Z}}(r)\,dr > \int_{w}^{v_{\mathcal{Z}}(x)} \lambda^x_{\mathcal{Z}}(r)\,dr$. Therefore, $\underline{G}_{x'}(w) < \underline{G}_x(w)$ for all $w \in [\underline{w}_{\mathcal{Z}}(x'), v_{\mathcal{Z}}(x)]$, and so $\underline{V}'_{\mathcal{Z}}(x') >_{FOSD} \underline{V}_{\mathcal{Z}}(x)$.

□

Proof of Proposition 1

Fix $p \in [0,1]$, fix $r, r' \in (0,1]$. Since $\succeq$ satisfies independence, we have $v_{\mathcal{Z}_r}(\ell_r(p)) = v_{\mathcal{Z}_r}(\ell_{r'}(p))$. Since τ has a CDF-complexity representation and therefore satisfies linearity, we have $\tau(\ell_r(p), (w, r)) = \tau(\ell'_r(p), (w, r'))$ for all $w \in \mathbb{R}$. Defining the valuation paradigm $((\ell_r(p), \mathcal{Z}_r), (\ell_{r'}(p), \mathcal{Z}_{r'}), \varphi)$, where $\varphi(v) = v$, Lemma 1 implies $V_{\mathcal{Z}_r}(\ell_r(p)) =_D V_{\mathcal{Z}_r}(\ell_{r'}(p))$, and so $\mathbb{E}[V_{\mathcal{Z}_r}(\ell_r(p))]$ is constant in r as desired. □

Proof of Proposition 2

Fix any $a, a' \in \mathbb{R}$. Define $x = (a, b, c; p, q, 1-p-q)$, $z_w = (a, b+m, c-w; p, q, 1-p-q)$, $\mathcal{Z} = \{z_w\}_{w \in [0, \overline{w}]}$ $x' = (a', b, c; p, q, 1-p-q)$, $z'_w = (a', b+m, c-w; p, q, 1-p-q)$, and $\mathcal{Z}' = \{z'_w\}_{w \in [0, \overline{w}]}$. Since $\succeq$ satisfies independence, we have $v_{\mathcal{Z}}(x) = v_{\mathcal{Z}'}(x')$. Since τ has a CDF-complexity representation and therefore satisfies linearity, we have $\tau(x, z_w) = \tau(x', z'_w)$ for all $w \in \mathbb{R}$. Defining the valuation paradigm $((x, \mathcal{Z}), (x', \mathcal{Z}'), \varphi)$, where $\varphi(v) = v$, Lemma 1 implies $V_{\mathcal{Z}}(x) =_D V_{\mathcal{Z}'}(x')$, and so $\mathbb{E}[V_{\mathcal{Z}}(x)]$ is constant in a as desired. □

Proof of Proposition 3

For $0 < m_L < m_H$, $p \in (0,1)$, define $\ell(1) = (m_L; 1)$, $\ell(p) = (m_L, m_H; p, 1-p)$, $z_w = (m_H; w)$

$\mathcal{Z} = \{z_w\}_{w \in [0,1]}$, and $\varphi(v) = 1 - 1/p \cdot (1-v)$. Let u be the Bernoulli utility function associated with $\succeq$; without loss, assume $u(0) = 0$. We have $v_{\mathcal{Z}}(\ell(1)) = u(m_L)/u(m_H)$ and $v_{\mathcal{Z}}(\ell(p)) = 1 - p\,(1 - u(m_L)/u(m_H))$, and so $\varphi(v_{\mathcal{Z}}(\ell(p)) = v_{\mathcal{Z}}(\ell(1))$.

Noting that $z_w = (1-p)z_1 + pz_{\varphi(w)}$ and $\ell(p) = (1-p)z_1 + p\ell(1)$, linearity of τ implies $\tau(\ell(1), z_{\varphi(w)}) = \tau(\ell(p), z_w)$ for all $w \in [0,1]$. Defining the valuation paradigm $((\ell(1), \mathcal{Z}), (\ell(p), \mathcal{Z}), \varphi)$, by Lemma 1, we have $V_{\mathcal{Z}}(\ell(1)) =_D \varphi(V_{\mathcal{Z}}(\ell(p)))$, which in turn implies $1 - \mathbb{E}[V_{\mathcal{Z}}(\ell(p))] = p\,(1 - \mathbb{E}[V_{\mathcal{Z}}(\ell(1))])$.

Proof of Proposition 4

Follows immediately from Lemma 1.

Proof of Proposition 5

Fix $n > 2$, and consider the multiple price list valuation (x, Z^n) of a DM with (u, τ^n, Q). Letting $\underline{w} = \underline{w}_{\mathcal{Z}}(x)$ and $\overline{w} = \overline{w}_{\mathcal{Z}}(x)$, define by $\Delta_n = (\overline{w} - \underline{w})/(n-1)$ the increment between each price, and for each $k \in \{1, ..., n\}$, define $z(k) = z_{w_k}$.

Fixing a realization of signals s, let $\overline{k}$ denote the index of the highest price for which the DM receives a signal that $z(k) \prec x$, and let $\underline{k}$ denote the index of the lowest price for which the DM receives a signal that $z(k) \succ x$; these prices are guaranteed to exist since $(x, \mathcal{Z})$ is bounded. Given s, the DM's valuation is given by $V(s) = (w_{\underline{k}} + w_{\overline{k}})/2$.

To show this, we first establish some basic notation and facts. Let $\pi(k)$ denote the event $u_{z(1)} < ... < u_{z(k-1)} < u_x < u_{z(k)} < ... < u_{z(n)}$. Under uniform priors, we have $\mathbb{E}[u_x|\pi(k)] \propto k$ and $\mathbb{E}[u_{z(j)}|\pi(k)] \propto j + \mathbb{1}(j \geq k)$. Since given s, the DM has uniform posteriors over the events $\pi(\underline{k}+1), ...\pi(\overline{k})$, the DM's posterior expected values are given by $\mathbb{E}[u_x|s] \propto (\underline{k} + \overline{k} + 1)/2$ $\mathbb{E}[u_{z(j)}|s] \propto j + (j - \underline{k})/(\overline{k} - \underline{k})$, for $\underline{k} \leq j \leq \overline{k}$. First, consider the case where $\overline{k} - \underline{k}$ is odd. For $R = (\underline{k} + \overline{k} - 1)/2$, we have $\mathbb{E}[u_{z(R)}|s] < \mathbb{E}[u_x|s] < \mathbb{E}[u_{z(R+1)}|s]$, and so the switching point is R, leading to the valuation $V(s) = (w_R + w_{R+1})/2 = (w_{\underline{k}} + w_{\overline{k}})/2$. Now consider the case where $\overline{k} - \underline{k}$ is even. In this case, for $R = \frac{\underline{k}+\overline{k}}{2}$, we have a tie: $E[u_x|s] = E[u_{z(R)}|s]$, leading to the valuation $V(s) = w_R = (w_{\underline{k}} + w_{\overline{k}})/2$.

Define the random variables $\overline{\delta}^n = w_{\overline{k}} - v_{\mathcal{Z}}(x)$ and $\underline{\delta}^n = v_{\mathcal{Z}}(x) - w_{\underline{k}}$; we have $V(s) = v_{\mathcal{Z}}(x) + \frac{1}{2}(\overline{\delta}^n - \underline{\delta}^n)$. Since $\overline{\delta}^n$ and $\underline{\delta}^n$ are independent by construction, all that remains is to show that $\overline{\delta}^n \xrightarrow{D} \overline{\delta}_{\mathcal{Z}}(x)$ and $\underline{\delta}^n \xrightarrow{D} \underline{\delta}_{\mathcal{Z}}(x)$. For $k \in \{1, ..., n\}$, define $\epsilon_k^n \sim_{iid} \text{Bernoulli}(\tau^n(x, z(k)))$. Define $S_n = \sum_{k=1}^n \epsilon_k^n \mathbb{1}(w_k \in [v_{\mathcal{Z}}(x), t])$, $p_k^n = \tau^n(x, z(k))\mathbb{1}(w_k \in [v_{\mathcal{Z}}(x), t])$, and $\Lambda_n = \sum_{k=1}^n p_k^n$. We have

$$\mathbb{P}(\overline{\delta}^n \leq t) = \begin{cases} 1 - \mathbb{P}\left(\sum_{k=1}^n S_n = 0\right) & t < \overline{w} - v_{\mathcal{Z}}(x) \\ 1 & \text{otherwise} \end{cases}$$

and Le Cam's theorem implies

$$\left|\mathbb{P}\left(\sum_{k=1}^{n} S_n = 0\right) - \exp(-\Lambda_n)\right| < 2\sum_{k=1}^{n}(p_k^n)^2$$

Defining $\lambda(t) = \frac{1}{\overline{w}-\underline{w}} \ln(1-\tau(x,z_t))$, note that $\tau^n(x,z(k)) = 1-\exp(-\lambda(w_k)\Delta_n)$. By assumption, $\lambda(t)$ is square-integrable on $[\underline{w},\overline{w}]$, and so $2\sum_{k=1}^{n}(p_k^n)^2 \to 0$ and $\Lambda_n \to \int_{v_{\mathcal{Z}}(x)}^{t} \lambda(r)\,dr = \int_0^t \lambda(v_{\mathcal{Z}}(x)+r)\,dr$, which in turn implies

$$\lim_{n\to\infty} \mathbb{P}(\overline{\delta}^n \leq t) = \begin{cases} 1-\exp\left(-\int_0^t \lambda(v_{\mathcal{Z}}(x)+r)\,dr\right) & t < \overline{w} - v_{\mathcal{Z}}(x) \\ 1 & \text{otherwise} \end{cases}$$

and so $\overline{\delta}^n \xrightarrow{D} \overline{\delta}_{\mathcal{Z}}(x)$. An analogous argument shows that $\underline{\delta}^n \xrightarrow{D} \underline{\delta}_{\mathcal{Z}}(x)$. □

Proof of Proposition 6

Since $x' >_{FOSD} x$, we have $\overline{w}_{\mathcal{Z}}(x) \leq \overline{w}_{\mathcal{Z}}(x')$ and $\underline{w}_{\mathcal{Z}}(x) \leq \underline{w}_{\mathcal{Z}}(x')$, and by assumption $\overline{w}_{\mathcal{Z}}(x') - \underline{w}_{\mathcal{Z}}(x') = \overline{w}_{\mathcal{Z}}(x) - \underline{w}_{\mathcal{Z}}(x)$. Furthermore, since $x' >_{FOSD} x$, monotonicity of $\succeq$ implies $v_{\mathcal{Z}}(x') > v_{\mathcal{Z}}(x)$, and monotonicity of τ implies that $\tau(x',z_w) < \tau(x,z_w)$ for $w \geq v_{\mathcal{Z}}(x')$ and $\tau(x',z_w) > \tau(x,z_w)$ for $w \leq v_{\mathcal{Z}}(x)$. By Lemma 2, we therefore have $V_{\mathcal{Z}}(x') >_{FOSD} V_{\mathcal{Z}}(x)$ as desired.

Proof of Proposition 7

Follows immediately from Lemma 1. □

Proof of Proposition 8

For $\gamma > 0$, define $G : [0,1] \to [0,1]$ by

$$\begin{aligned} G(x) = x + \frac{1}{2}\int_x^1 \exp\left(\gamma \int_x^w \log\left(\frac{2x(1-t)}{x+t(1-2x)}\right) dt\right) dw \\ -\frac{1}{2}\int_0^x \exp\left(\gamma \int_w^x \log\left(\frac{2(1-x)t}{x+t(1-2x)}\right) dt\right) dw. \end{aligned}$$

It will be useful to observe the following properties of G.

Lemma 3. *G has the following properties:*

i) If $x \leq 1/2$, then $G(x) \geq x$. If $x < 1/2$, then $G(x) > x$.

ii) *If $x \geq 1/2$, then $G(x) \geq 1/2$.*

iii) $G_1(1-x) = 1 - G_1(x)$.

Proof. Define $l(t) = \log\left(\frac{t}{1-t}\right)$, $\phi(s) = \log\left(\frac{1+e^s}{2}\right)$, and $C_x(t) = \phi(|l(t) - l(x)|)$. We have

$$G_\alpha(x) = x + \frac{1}{2}R(x) - \frac{1}{2}L(x)$$

where $R(x) = \int_x^1 \exp\left(-\gamma \int_x^w C_x(t)dt\right)dw$ and $L(x) = \int_0^x \exp\left(-\gamma \int_w^x C_x(t)dt\right)dw$. Fix $\alpha \geq 1$. We prove the first three parts in order.

For part (i), assume $x \leq 1/2$. It is enough to show $R(x) \geq L(x)$. For $s \in [0,x]$, note that $l(x+s) + l(x-s)$ is weakly decreasing in s. This implies $l(x+s) - l(x) \leq l(x) - l(x-s)$. Since ϕ is increasing, we have $C_x(x+s) \leq C_x(x-s)$ for $s \in [0,x]$, and so

$$\int_0^x \exp\left(-\gamma \int_0^s C_x(x+v)dv\right)ds \geq \int_0^x \exp\left(-\gamma \int_0^s C_x(x-v)dv\right)ds$$

Since $L(x) = \int_0^x \exp\left(-\gamma \int_0^s C_x(x-v)dv\right)ds$ and $R(x) \geq \int_0^x \exp\left(-\gamma \int_0^s C_x(x+v)dv\right)ds$ whenever $x \leq 1/2$, with a strict inequality whenever $x < 1/2$, the desired inequalities for part (i) obtain.

For part ii), assume $x \geq 1/2$. For $t \in [x,1]$, note that $l(t) + l(1-t)$ is weakly decreasing in t. Therefore, we have $l(t)+l(1-t) \leq l(x)+l(1-x) \leq 2l(x)$. Hence $l(t)-l(x) \leq l(x)-l(1-t)$, and therefore $C_x(t) \leq C_x(1-t)$. This implies that for all $w \in [x,1]$,

$$\int_x^w C_x(t)dt \leq \int_x^w C_x(1-t)dt = \int_{1-w}^{1-x} C_x(t)dt \leq \int_{1-w}^{x} C_x(t)dt.$$

It follows that

$$\int_x^1 \left(1 - \exp\left(-\gamma \int_x^w C_x(t)dt\right)\right)dw \leq \int_0^x \left(1 - \exp\left(-\gamma \int_w^x C_x(t)dt\right)\right)dw$$

and so $(1-x) - R(x) \leq x - L(x) \implies G_\alpha(x) - 1/2 \geq 0$. This proves part ii).

To show part iii), let $A_x(t) = \frac{2x(1-t)}{t+x(1-2t)}$, $B_x(t) = \frac{2(1-x)t}{t+x(1-2t)}$. We have $A_{1-x}(1-t) = B_x(t)$ and $B_{1-x}(1-t) = A_x(t)$, and so a change of variables $t \mapsto 1-t$ and $w \mapsto 1-w$ in the two integral terms in the definition of $G_1(1-x)$ yields $G_1(1-x) = 1 - G_1(x)$. □

It will also be useful to define the average certainty equivalents $\pi(p) = \mathbb{E}[V_{\mathcal{Z}}((1;p))]$, for $\mathcal{Z} = \{(w;1)\}_{w \in \mathbb{R}}$, and the average probability equivalents $\psi(c) = \mathbb{E}[V_{\mathcal{Z}'}((c;1))]$, for $\mathcal{Z}' = \{(1;w)\}_{w \in [0,1]}$.

Lemma 4. *Suppose $\succeq$ has an EU representation $u(x) = x^\alpha$ and $\tau(\cdot,\cdot)$ has a CDF-complexity representation (u,H) with $H(r) = 1-(1-r)^\gamma$. For all $p,c \in (0,1)$, $\pi(p)$ is increasing in α and $\psi(c)$ is decreasing in α.*

Proof. Fix $p \in (0,1)$. We first show that $\pi(p)$ is increasing in α. Let $x = (1;p)$ and $\mathcal{Z} = \{(w;1)\}_{w\in\mathbb{R}}$. Let $V(\alpha)$ and $v(\alpha) = p^{1/\alpha}$ denote the observed and true valuations, $V_{\mathcal{Z}}(x)$ and $v_{\mathcal{Z}}(x)$, as a function of α, and let $\tau^\alpha(\cdot,\cdot)$ denote the comparability function given α. Fix $\alpha' > \alpha$; it suffices to show that $V(\alpha') >_{FOSD} V(\alpha)$. Note that $v(\alpha') > v(\alpha)$. Since for $w > v(\alpha)$, $\tau^\alpha(x,z_w) = \frac{w^\alpha - p}{p+w^\alpha(1-2p)}$ is decreasing in α, we have $\tau^{\alpha'}(x,z_w) < \tau^\alpha(x,z_w)$ for all $w > v(\alpha')$. Similarly, since for $w < v(\alpha')$, $\tau^{\alpha'}(x,z_w) = \frac{p-w^{\alpha'}}{p+w^{\alpha'}(1-2p)}$ is increasing in α', we have $\tau^{\alpha'}(x,z_w) > \tau^\alpha(x,z_w)$ for all $w < v(\alpha)$. By Lemma 2, we have $V(\alpha') >_{FOSD} V(\alpha)$ as desired.

Fix $c \in (0,1)$. We first show that $\psi(c)$ is decreasing in α. Let $y = (c;1)$ and $\mathcal{Z}' = \{(1;w)\}_{w\in[0,1]}$. Let $V(\alpha)$ and $v(\alpha) = c^\alpha$ denote the observed and true valuations, $V_{\mathcal{Z}'}(y)$ and $v_{\mathcal{Z}'}(y)$, as a function of α, and let $\tau^\alpha(\cdot,\cdot)$ denote the comparability function given α. Fix $\alpha' > \alpha$; it suffices to show that $V(\alpha) >_{FOSD} V(\alpha')$. Note that $v(\alpha) > v(\alpha')$. Since for $w > v(\alpha')$, $\tau^{\alpha'}(x,z_w) = \frac{w-c^{\alpha'}}{w+c^{\alpha'}(1-2w)}$ is increasing in α', we have $\tau^{\alpha'}(x,z_w) > \tau^\alpha(x,z_w)$ for all $w > v(\alpha)$. Similarly, since for $w < v(\alpha)$, $\tau^\alpha(x,z_w) = \frac{c^\alpha - w}{p+w^\alpha(1-2p)}$ is decreasing in α, we have $\tau^{\alpha'}(x,z_w) < \tau^\alpha(x,z_w)$ for all $w < v(\alpha')$. By Lemma 2, we have $V(\alpha') >_{FOSD} V(\alpha)$ as desired. □

Lemma 1 yields the identities

$$\begin{aligned} \mathbb{E}[V_{M\text{-}a}(\ell_R)] &= 30 \cdot \pi(p), & \mathbb{E}[V_{M\text{-}a}(\ell_S)] &= 30p \\ \mathbb{E}[V_{P\text{-}a}(\ell_R)] &= 0.2p, & \mathbb{E}[V_{P\text{-}a}(\ell_S)] &= 0.2 \cdot \psi(p) \\ \mathbb{E}[V_{P\text{-}b}(\ell_R)] &= 0.2p \cdot \psi(0.75), & \mathbb{E}[V_{P\text{-}b}(\ell_S)] &= 0.2 \cdot \psi(0.75p), \end{aligned}$$

and for $\alpha = 1$, we have $G(x) = \pi(x) = \psi(x)$ by construction.

Begin with the proof of (a), and fix $p \in \{0.1, 0.2\}$. For $\alpha = 1$, Lemmas 3-i) and 3-iii) jointly imply $G(p) > p$, which, since $G(p) = \pi(p) = \psi(p)$ for $\alpha = 1$, implies $\mathbb{E}[V_{M\text{-}a}(\ell_R)] > \mathbb{E}[V_{M\text{-}a}(\ell_S)] = 30p$ and $0.2p = \mathbb{E}[V_{P\text{-}a}(\ell_R)] < \mathbb{E}[V_{P\text{-}a}(\ell_S)]$. Now suppose that $\mathbb{E}[V_{M\text{-}a}(\ell_R)] \leq \mathbb{E}[V_{M\text{-}a}(\ell_S)]$. Since $\mathbb{E}[V_{M\text{-}a}(\ell_R)]$ is increasing in α by Lemma 4, it must be the case that $\alpha < 1$. However, since Lemma 4 also implies that $\mathbb{E}[V_{P\text{-}a}(\ell_S)]$ is decreasing in α, it must be the case that $\mathbb{E}[V_{P\text{-}a}(\ell_R)] < \mathbb{E}[V_{P\text{-}a}(\ell_S)]$, as desired.

Now, fix $p \in \{0.8, 0.9\}$. For $\alpha = 1$, Lemma 3-i) implies $G(p) < p$, which, since $G(p) = \pi(p) = \psi(p)$ for $\alpha = 1$, implies $\mathbb{E}[V_{M\text{-}a}(\ell_R)] < \mathbb{E}[V_{M\text{-}a}(\ell_S)] = 30p$ and $0.2p = \mathbb{E}[V_{P\text{-}a}(\ell_R)] > \mathbb{E}[V_{P\text{-}a}(\ell_S)]$. Now suppose that $\mathbb{E}[V_{M\text{-}a}(\ell_R)] \geq \mathbb{E}[V_{M\text{-}a}(\ell_S)]$. Since $\mathbb{E}[V_{M\text{-}a}(\ell_R)]$ is increasing in α by Lemma 4, it must be the case that $\alpha > 1$. However, since Lemma 4 also implies that $\mathbb{E}[V_{P\text{-}a}(\ell_S)]$ is decreasing in α, it must be the case that $\mathbb{E}[V_{P\text{-}a}(\ell_R)] > \mathbb{E}[V_{P\text{-}a}(\ell_S)]$, as desired.

For the proof of (b), fix $p \in \{0.1, 0.2\}$, and suppose $\mathbb{E}[V_{M\text{-}a}(\ell_R)] \leq \mathbb{E}[V_{M\text{-}a}(\ell_S)]$. By the argument above, it must be the case that $\alpha < 1$. It suffices to show that $\psi(0.75p) > p\psi(0.75)$ for such α. Notice that by construction, we have $\psi(p) = G(p^\alpha)$, and so it suffices to show that $G((0.75p)^\alpha) > pG(0.75^\alpha)$. Since $0.75^\alpha > 1/2$, Lemmas 3-i) and 3-iii) jointly imply $G(0.75^\alpha) < 0.75^\alpha$. If $(0.75p)^\alpha < 1/2$, Lemma 3-i) yields $G((0.75p)^\alpha) > (0.75p)^\alpha > p0.75^\alpha > pG(0.75^\alpha)$ as desired. If instead $(0.75p)^\alpha \leq 1/2$, Lemma 3-ii) yields $G(0.75^\alpha) \geq 1/2 > pG(0.75^\alpha)$ as desired, since $p \leq 1/2$. □

Proof of Proposition 9

The proof of Proposition 8 establishes the following: 1) for $p \in \{0.1, 0.2\}$, $\mathbb{E}[V_{M\text{-}a}(\ell_R)] \leq \mathbb{E}[V_{M\text{-}a}(\ell_S)]$ implies $\alpha < 1$ and $\mathbb{E}[V_{P\text{-}a}(\ell_R)] \geq \mathbb{E}[V_{P\text{-}a}(\ell_S)]$ implies $\alpha > 1$, and 2) for $p \in \{0.8, 0.9\}$, $\mathbb{E}[V_{M\text{-}a}(\ell_R)] \geq \mathbb{E}[V_{M\text{-}a}(\ell_S)]$ implies $\alpha > 1$ and $\mathbb{E}[V_{P\text{-}a}(\ell_R)] \leq \mathbb{E}[V_{P\text{-}a}(\ell_S)]$ implies $\alpha < 1$.

To show (a), fix $p \in \{0.8, 0.9\}$ and suppose $\mathbb{E}[V_{M\text{-}a}(\ell_R)] \geq \mathbb{E}[V_{M\text{-}a}(\ell_S)]$. The above implies $\alpha > 1$, which by assumption implies $\mathbb{E}[V_{M\text{-}b}(\ell_R)] > \mathbb{E}[V_{M\text{-}b}(\ell_S)]$ as desired. To show (b), fix $p \in \{0.1, 0.2\}$ and suppose $\mathbb{E}[V_{P\text{-}a}(\ell_R)] \geq \mathbb{E}[V_{P\text{-}a}(\ell_S)]$. The above implies $\alpha > 1$, which by assumption implies $\mathbb{E}[V_{M\text{-}b}(\ell_R)] > \mathbb{E}[V_{M\text{-}b}(\ell_S)]$ as desired. To show (c), fix $p \in \{0.8, 0.9\}$ and suppose $\mathbb{E}[V_{P\text{-}a}(\ell_R)] \leq \mathbb{E}[V_{P\text{-}a}(\ell_S)]$. The above implies $\alpha < 1$, which by assumption implies $\mathbb{E}[V_{P\text{-}b}(\ell_R)] < \mathbb{E}[V_{P\text{-}b}(\ell_S)]$ as desired. Finally, to show (d), suppose $p \in \{0.1, 0.2, 0.8, 0.9\}$ and suppose $\mathbb{E}[V_{M\text{-}b}(\ell_R)] \leq \mathbb{E}[V_{M\text{-}b}(\ell_S)]$. By assumption, it must be the case that $\alpha < 1$, which by assumption implies $\mathbb{E}[V_{P\text{-}b}(\ell_R)] < \mathbb{E}[V_{P\text{-}b}(\ell_S)]$ as desired.

C Experimental Design: Details

C.1 Protocol and Design Details

Multiple Price Lists. All valuations are elicited using multiple price lists with evenly spaced prices, where the endpoints of the list correspond to the natural dominance boundaries of the valuation task. Price list increments were chosen to keep price list length comparable across valuation formats, while maintaining simple increments.

- M-a valuations: ℓ_R is valued against $Z = (z^1, ..., z^n)$, where $z^1 = (0; 0.2)$, $z^n = (30; 0.2)$, and $n = 21$ for all p.
- M-b valuations: ℓ_R is valued against $Z = (z^1, ..., z^n)$, where $z^1 = (0; 0.5)$, $z^n = (30; 0.5)$, and $n = 21$ for all p. ℓ_S is valued against $Z = (z^1, ..., z^n)$, where $z^1 = (0; 0.5)$, $z^n = (30p; 0.5)$, and $n = 21, 19, 21, 17, 18, 19$ for $p = 0.1, 0.2, 0.8, 0.9$, respectively.

- P-a valuations: ℓ_S is valued against $Z = (z^1, ..., z^n)$, where $z^1 = (30; 0)$, $z^n = (30; 0.2)$, and $n = 21$ for all p.
- P-b valuations: ℓ_S is valued against $Z = (z^1, ..., z^n)$, where $z^1 = (40; 0)$, $z^n = (40; 0.2)$, and $n = 21$ for all p. ℓ_R is valued against $Z = (z^1, ..., z^n)$, where $z^1 = (40; 0)$, $z^n = (40; 0.2p)$, and $n = 21, 21, 21, 17, 18, 19$ for all $p = 0.1, 0.2, 0.8, 0.9$, respectively.

C.2 Experimental Interface

Instructions, Money Equivalents Block

Part 1 Instructions

Please read these instructions carefully. There will be comprehension checks. If you fail these checks, you will be excluded from the study and you will not receive the completion payment.

Your task

There are 13 rounds in this part of the study. In each round, you will make a series of choices between a fixed option ("Option A") and a range of alternatives ("Option B"). Each of these options are **lotteries** that pay out an amount of money with some percentage chance, and pay out nothing otherwise.

Here is an example of a choice list similar to those you will complete in each round, where each row in this list is a separate choice.

Choice	Option A (fixed)		Option B (varies)
#1	$8.00 with **30% chance**	or	$0.00 with **40% chance**
#2	$8.00 with **30% chance**	or	$1.00 with **40% chance**
#3	$8.00 with **30% chance**	or	$2.00 with **40% chance**
#4	$8.00 with **30% chance**	or	$3.00 with **40% chance**
#5	$8.00 with **30% chance**	or	$4.00 with **40% chance**
#6	$8.00 with **30% chance**	or	$5.00 with **40% chance**
#7	$8.00 with **30% chance**	or	$6.00 with **40% chance**
#8	$8.00 with **30% chance**	or	$7.00 with **40% chance**
#9	$8.00 with **30% chance**	or	$8.00 with **40% chance**

- In each choice list, **Option B** (right side) will have a set chance of paying out some amount of money. This payment amount **increases** as you go down the list. In this example, Option B pays out an amount between $0 and $8 with a 40% chance, and pays $0 otherwise.
- **Option A** (left side) will be the **same in all rows**. Here, Option A pays out $8 with a 30% chance, and pays $0 otherwise.
- To make a choice, just click on the option you prefer for each choice (i.e. for each row), and the computer will highlight your choice.
- Since Option B pays out progressively more as you go down the list, **we assume you will choose Option A at first, and at some point may switch to choosing Option B**. You can click on your choice in the row that you would "switch" from A to B, and we will automatically fill out the rest of the list for you. We do this by selecting Option A in all rows above and Option B in all rows below your selected row. **Try this on the list above!**

Your bonus payment

Your decisions in this part of the study may affect your bonus. If a round in this part is selected for payment, the computer will randomly select one of your individual choices. The computer will then play out the lottery you chose, and pay you the outcome of the lottery.

- For example, if choice #6 in the list below was selected for payment, you'd have a 30% chance of earning $8 and a 70% chance of earning $0.
- If instead choice #7 was selected for payment, you'd have a 40% chance of earning $6 and a 60% chance of earning $0.

Choice	Option A (fixed)		Option B (varies)
#1	$8.00 with **30% chance**	or	$0.00 with **40% chance**
#2	$8.00 with **30% chance**	or	$1.00 with **40% chance**
#3	$8.00 with **30% chance**	or	$2.00 with **40% chance**
#4	$8.00 with **30% chance**	or	$3.00 with **40% chance**
#5	$8.00 with **30% chance**	or	$4.00 with **40% chance**
#6	$8.00 with **30% chance**	or	$5.00 with **40% chance**
#7	$8.00 with **30% chance**	or	$6.00 with **40% chance**
#8	$8.00 with **30% chance**	or	$7.00 with **40% chance**
#9	$8.00 with **30% chance**	or	$8.00 with **40% chance**

These rounds are completely independent from one another. When a round is selected to determine your bonus, only your choices in that one round will determine your bonus. This means that it is in your best interest to **choose the option you actually prefer in each case.**

Attention check round

One of the 13 rounds in this part of the study will be an **attention check**.

- This attention check round should be very easy to complete if you are paying attention. This is because every row in this particular choice list will contain a **payoff-maximizing option**: an option thay pays out *more money* with a *higher chance* than the other.
- To pass this attention check, you must **select the payoff-maximizing option in every row** of the choice list. Note that this is also exactly what you would do if you were trying to maximize your bonus.

There will be three attention checks throughout the entire study, inlcuding this one. **If you fail two or more attention checks, you will not be allowed to continue with the study**. Therefore, to complete the study, it is important to pay attention to the task.

Comprehension Checks, Money Equivalents Block

Comprehension check

To verify your understanding of the instructions, please answer the comprehension questions below. If you get one or more of them wrong twice in a row, you will not be allowed to participate in the study and earn a completion payment. In each question, exactly one response option is correct.

You can review the instructions here.

1. Which one of the following statements is true?

- I should complete the choice lists by thinking carefully about the row in which I would like to switch from preferring Option A (the fixed option) to preferring Option B (the option with the varying payment amount).
- I should always switch in the last row of the choice list, because this increases the chance that Option A (the fixed option) will determine my bonus.
- I should always switch in the first row of the choice list, because this increases the chance that Option B (the option with the varying payment amount) will determine my bonus.

2. Suppose a round is selected to determine your bonus. Which one of the following statements is true?

- The computer will randomly select one of the rows of that round, and I will receive the option I selected in that row.
- I will receive the option I selected in every row of that round.
- I will receive the option I selected in the first row of that round.
- I will receive the option I selected in each row of that round.

3. Imagine a participant in this study makes the following choices. According to these choices, which of the following statements about this participant is false?

Choice	Option A (fixed)		Option B (varies)
#1	$8.00 with 30% chance	or	$0.00 with 40% chance
#2	$8.00 with 30% chance	or	$1.00 with 40% chance
#3	$8.00 with 30% chance	or	$2.00 with 40% chance
#4	$8.00 with 30% chance	or	$3.00 with 40% chance
#5	$8.00 with 30% chance	or	$4.00 with 40% chance
#6	$8.00 with 30% chance	or	$5.00 with 40% chance
#7	$8.00 with 30% chance	or	$6.00 with 40% chance
#8	$8.00 with 30% chance	or	$7.00 with 40% chance
#9	$8.00 with 30% chance	or	$8.00 with 40% chance

The participant prefers getting Option A over getting $5 with 40% chance

The participant prefers getting Option A over getting $3 with 40% chance

The participant prefers getting $5 with 40% chance over getting Option A

The participant prefers getting $6 with 40% chance over getting Option A

4. Suppose the following choice list appeared in an attention check round. To pass the attention check, how should you fill out this list?

Choice	Option A (varies)		Option B (fixed)
#1	$8.00 with 20% chance	or	$0.00 with 15% chance
#2	$8.00 with 20% chance	or	$1.00 with 15% chance
#3	$8.00 with 20% chance	or	$2.00 with 15% chance
#4	$8.00 with 20% chance	or	$3.00 with 15% chance
#5	$8.00 with 20% chance	or	$4.00 with 15% chance
#6	$8.00 with 20% chance	or	$5.00 with 15% chance
#7	$8.00 with 20% chance	or	$6.00 with 15% chance
#8	$8.00 with 20% chance	or	$7.00 with 15% chance
#9	$8.00 with 20% chance	or	$8.00 with 15% chance

Task Screen, Money Equivalents Block

Please select the option you prefer in each row.

Choice	Option A (fixed)		Option B (varies)
#1	$30.00 with 18% chance	or	$0.00 with 50% chance
#2	$30.00 with 18% chance	or	$1.50 with 50% chance
#3	$30.00 with 18% chance	or	$3.00 with 50% chance
#4	$30.00 with 18% chance	or	$4.50 with 50% chance
#5	$30.00 with 18% chance	or	$6.00 with 50% chance
#6	$30.00 with 18% chance	or	$7.50 with 50% chance
#7	$30.00 with 18% chance	or	$9.00 with 50% chance
#8	$30.00 with 18% chance	or	$10.50 with 50% chance
#9	$30.00 with 18% chance	or	$12.00 with 50% chance
#10	$30.00 with 18% chance	or	$13.50 with 50% chance
#11	$30.00 with 18% chance	or	$15.00 with 50% chance
#12	$30.00 with 18% chance	or	$16.50 with 50% chance
#13	$30.00 with 18% chance	or	$18.00 with 50% chance
#14	$30.00 with 18% chance	or	$19.50 with 50% chance
#15	$30.00 with 18% chance	or	$21.00 with 50% chance
#16	$30.00 with 18% chance	or	$22.50 with 50% chance
#17	$30.00 with 18% chance	or	$24.00 with 50% chance
#18	$30.00 with 18% chance	or	$25.50 with 50% chance
#19	$30.00 with 18% chance	or	$27.00 with 50% chance
#20	$30.00 with 18% chance	or	$28.50 with 50% chance
#21	$30.00 with 18% chance	or	$30.00 with 50% chance

Submit Choices

Instructions, Probability Equivalents Block

Part 1 Instructions

Please read these instructions carefully. There will be comprehension checks. If you fail these checks, you will be excluded from the study and you will not receive the completion payment.

Your task

There are 13 rounds in this part of the study. In each round, you will make a series of choices between a fixed option ("Option A") and a range of alternatives ("Option B"). Each of these options are **lotteries** that pay out an amount of money with some percentage chance, and pay out nothing otherwise.

Here is an example of a choice list similar to those you will complete in each round, where each row in this list is a separate choice.

Choice	Option A (fixed)		Option B (varies)
#1	$3.00 with **40% chance**	or	$8.00 with **0% chance**
#2	$3.00 with **40% chance**	or	$8.00 with **5% chance**
#3	$3.00 with **40% chance**	or	$8.00 with **10% chance**
#4	$3.00 with **40% chance**	or	$8.00 with **15% chance**
#5	$3.00 with **40% chance**	or	$8.00 with **20% chance**
#6	$3.00 with **40% chance**	or	$8.00 with **25% chance**
#7	$3.00 with **40% chance**	or	$8.00 with **30% chance**
#8	$3.00 with **40% chance**	or	$8.00 with **35% chance**
#9	$3.00 with **40% chance**	or	$8.00 with **40% chance**

- In each choice list, **Option B** (right side) will pay out a set amount of money with some percent chance. This percent chance **increases** as you go down the list. In this example, Option B pays out $8 with a percent chance ranging from 0% to 40%, and pays out $0 otherwise.
- **Option A** (left side) will be the **same in all rows**. Here, Option A pays out $3 with a 40% chance, and pays $0 otherwise.
- To make a choice, just click on the option you prefer for each choice (i.e. for each row), and the computer will highlight your choice.
- Since Option B has a progressively higher chance of paying out as you go down the list, **we assume you will choose Option A at first, and at some point may switch to choosing Option B**. You can click on your choice in the row that you would "switch" from A to B, and we will automatically fill out the rest of the list for you. We do this by selecting Option A in all rows above and Option B in all rows below your selected row. **Try this on the list above!**

Your bonus payment

Your decisions in this part of the study may affect your bonus. If a round in this part is selected for payment, the computer will randomly select one of your individual choices. The computer will then play out the lottery you chose, and pay you the outcome of the lottery.

- For example, if choice #4 in the list below was selected for payment, you'd have a 40% chance of earning $3 and a 60% chance of earning $0.
- If instead choice #5 was selected for payment, you'd have a 20% chance of earning $8 and a 80% chance of earning $0.

Choice	Option A (fixed)		Option B (varies)
#1	$3.00 with **40% chance**	or	$8.00 with **0% chance**
#2	$3.00 with **40% chance**	or	$8.00 with **5% chance**
#3	$3.00 with **40% chance**	or	$8.00 with **10% chance**
#4	$3.00 with **40% chance**	or	$8.00 with **15% chance**
#5	$3.00 with **40% chance**	or	$8.00 with **20% chance**
#6	$3.00 with **40% chance**	or	$8.00 with **25% chance**
#7	$3.00 with **40% chance**	or	$8.00 with **30% chance**
#8	$3.00 with **40% chance**	or	$8.00 with **35% chance**
#9	$3.00 with **40% chance**	or	$8.00 with **40% chance**

These rounds are completely independent from one another. When a round is selected to determine your bonus, only your choices in that one round will determine your bonus. This means that it is in your best interest to **choose the option you actually prefer in each case.**

Attention check round

One of the 13 rounds in this part of the study will be an **attention check**.

- This attention check round should be very easy to complete if you are paying attention. This is because every row in this particular choice list will contain a **payoff-maximizing option**: an option thay pays out *more money* with a *higher chance* than the other.
- To pass this attention check, you must **select the payoff-maximizing option in every row** of the choice list. Note that this is also exactly what you would do if you were trying to maximize your bonus.

There will be three attention checks throughout the entire study, inlcuding this one. **If you fail two or more attention checks, you will not be allowed to continue with the study**. Therefore, to complete the study, it is important to pay attention to the task.

Comprehension Checks, Probability Equivalents Block

Comprehension check

To verify your understanding of the instructions, please answer the comprehension questions below. If you get one or more of them wrong twice in a row, you will not be allowed to participate in the study and earn a completion payment. In each question, exactly one response option is correct.

You can review the instructions here.

1. Which one of the following statements is true?

- I should always switch in the first row of the choice list, because this increases the chance that Option B (the option with the varying chance of payment) will determine my bonus.
- I should always switch in the last row of the choice list, because this increases the chance that Option A (the fixed option) will determine my bonus.
- I should complete the choice lists by thinking carefully about the row in which I would like to switch from preferring Option A (the fixed option) to preferring Option B (the option with the varying chance of payment).

2. Suppose a round is selected to determine your bonus. Which one of the following statements is true?

- I will receive the option I selected in each row of that round.
- I will receive the option I selected in every row of that round.
- I will receive the option I selected in the first row of that round.
- The computer will randomly select one of the rows of that round, and I will receive the option I selected in that row.

3. Imagine a participant in this study makes the following choices. According to these choices, which of the following statements about this participant is false?

Choice	Option A (fixed)		Option B (varies)
#1	$3.00 with **40% chance**	or	$8.00 with **0% chance**
#2	$3.00 with **40% chance**	or	$8.00 with **5% chance**
#3	$3.00 with **40% chance**	or	$8.00 with **10% chance**
#4	$3.00 with **40% chance**	or	$8.00 with **15% chance**
#5	$3.00 with **40% chance**	or	$8.00 with **20% chance**
#6	$3.00 with **40% chance**	or	$8.00 with **25% chance**
#7	$3.00 with **40% chance**	or	$8.00 with **30% chance**
#8	$3.00 with **40% chance**	or	$8.00 with **35% chance**
#9	$3.00 with **40% chance**	or	$8.00 with **40% chance**

- The participant prefers getting Option A over getting $8 with 20% chance
- The participant prefers getting $8 with 20% chance over getting Option A
- The participant prefers getting Option A over getting $8 with 15% chance
- The participant prefers getting $8 with 30% chance over getting Option A

4. Suppose the following choice list appeared in an attention check round. To pass the attention check, how should you fill out this list?

Choice	Option A (varies)		Option B (fixed)
#1	$15.00 with **40% chance**	or	$12.00 with **0% chance**
#2	$15.00 with **40% chance**	or	$12.00 with **5% chance**
#3	$15.00 with **40% chance**	or	$12.00 with **10% chance**
#4	$15.00 with **40% chance**	or	$12.00 with **15% chance**
#5	$15.00 with **40% chance**	or	$12.00 with **20% chance**
#6	$15.00 with **40% chance**	or	$12.00 with **25% chance**
#7	$15.00 with **40% chance**	or	$12.00 with **30% chance**
#8	$15.00 with **40% chance**	or	$12.00 with **35% chance**
#9	$15.00 with **40% chance**	or	$12.00 with **40% chance**

Task Screen, Probability Equivalents Block

Please select the option you prefer in each row.

Choice	Option A (fixed)		Option B (varies)
#1	$30.00 with **18% chance**	or	$40.00 with **0% chance**
#2	$30.00 with **18% chance**	or	$40.00 with **1% chance**
#3	$30.00 with **18% chance**	or	$40.00 with **2% chance**
#4	$30.00 with **18% chance**	or	$40.00 with **3% chance**
#5	$30.00 with **18% chance**	or	$40.00 with **4% chance**
#6	$30.00 with **18% chance**	or	$40.00 with **5% chance**
#7	$30.00 with **18% chance**	or	$40.00 with **6% chance**
#8	$30.00 with **18% chance**	or	$40.00 with **7% chance**
#9	$30.00 with **18% chance**	or	$40.00 with **8% chance**
#10	$30.00 with **18% chance**	or	$40.00 with **9% chance**
#11	$30.00 with **18% chance**	or	$40.00 with **10% chance**
#12	$30.00 with **18% chance**	or	$40.00 with **11% chance**
#13	$30.00 with **18% chance**	or	$40.00 with **12% chance**
#14	$30.00 with **18% chance**	or	$40.00 with **13% chance**
#15	$30.00 with **18% chance**	or	$40.00 with **14% chance**
#16	$30.00 with **18% chance**	or	$40.00 with **15% chance**
#17	$30.00 with **18% chance**	or	$40.00 with **16% chance**
#18	$30.00 with **18% chance**	or	$40.00 with **17% chance**
#19	$30.00 with **18% chance**	or	$40.00 with **18% chance**

Submit Choices