No. 97

UNIVERSITY OF COLOGNE

WORKING PAPER SERIES IN ECONOMICS

LOCAL THINKING AND SKEWNESS PREFERENCES

MARKUS DERTWINKEL-KALT
MATS K¥ STER

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

Local Thinking and Skewness Preferences*

Markus Dertwinkel-Kalt[1,†] Mats Köster[2,‡]

[1]University of Cologne
[2]Düsseldorf Institute for Competition Economics (DICE)

First version: February 2017
This version: July 2017

Abstract

We show that continuous models of stimulus-driven attention can account for skewness-related puzzles in decision-making under risk. First, we delineate that these models provide a well-defined theory of choice under risk. We therefore prove that in continuous—in contrast to discrete—models of stimulus-driven attention each lottery has a unique certainty equivalent that is monotonic in probabilities (i.e., it monotonically increases if probability mass is shifted to more favorable outcomes). Second, we show that whether an agent seeks or avoids a specific risk depends on the skewness of the underlying probability distribution. Since unlikely, but outstanding payoffs attract attention, an agent exhibits a preference for right-skewed and an aversion toward left-skewed risks. While cumulative prospect theory can also account for such skewness preferences, it yields implausible predictions on their magnitude. We show that these extreme implications can be ruled out for continuous models of stimulus-driven attention.

JEL-Classification: D81.

Keywords: Stimulus-Driven Attention; Salience Theory; Focusing; Certainty Equivalent; Monotonicity; Skewness Preferences.

*We thank Paul Heidhues and Daniel Wiesen for very helpful comments and suggestions. We also thank seminar audiences at HHU Düsseldorf, Goethe-University Frankfurt, LMU München, and University of Hamburg as well as participants at the ECORES Summer School 2017 (Louvain-la-Neuve) and the briq Summer School 2017 (Bonn). Mats Köster gratefully acknowledges financial support through the Graduate Programme Competition Economics (GRK 1974) of the German Science Foundation (DFG).

†Email: markus.dertwinkel-kalt@uni-koeln.de
‡Email: mats.koester@hhu.de

1 Introduction

Few individuals are globally risk-averse or risk-seeking. Instead, many individuals buy insurance (i.e., behave as if risk-averse) and gamble in casinos (i.e., behave as if risk-seeking) at the same time. Whether an agent seeks or avoids a specific risk depends on the skewness of the underlying probability distribution. Typically, agents insure against large potential losses that rarely occur (e.g., Sydnor, 2010; Barseghyan *et al.*, 2013). For example, natural disasters belong to this group of left-skewed risks. At the same time, many individuals seek right-skewed risks such as casino gambling according to which a large gain is realized with a very small probability (e.g., Golec and Tamarkin, 1998; Garrett and Sobel, 1999; Forrest *et al.*, 2002). The observation that agents tend to seek right-skewed and avoid left-skewed risks is referred to as *skewness preferences*.

A compelling explanation for skewness preferences is still missing. As expected utility theory (EUT) implies a valuation for risky options that is linear in probabilities, it typically predicts either risk-averse *or* risk-seeking behavior. In particular, for all commonly used utility functions, EUT cannot account for skewness preferences. In order to match experimental and empirical evidence, cumulative prospect theory (CPT; Tversky and Kahneman, 1992) has proposed a non-linear probability weighting. As a CPT agent overweights small probabilities by assumption, she exhibits a preference for right-skewed and an aversion toward left-skewed risks. This mechanism, however, does not offer any psychologically sound explanation for why skewness matters. In addition, cumulative prospect theory makes implausible predictions on the magnitude of skewness preferences (e.g., Rieger and Wang, 2006; Azevedo and Gottlieb, 2012; Ebert and Strack, 2015, 2016). Altogether, neither expected utility theory nor cumulative prospect theory convincingly address the role of skewness in choice under risk.

Models of stimulus-driven attention offer a more intuitive explanation for skewness preferences. According to these models, individuals are *local thinkers* whose attention is automatically directed toward certain outstanding choice features while less attention-grabbing aspects tend to be neglected.[1] Similar to cumulative prospect theory, these approaches incorporate non-linear probability weighting, but the distortion of a probability weight is endogenously determined by the relative size of the corresponding payoff. Probabilities of outstanding outcomes are inflated, while probabilities of less attention-grabbing outcomes are underweighted. In a typical lottery game, for instance, the large jackpot stands out relative to the rather low price of the lottery ticket, thereby attracting a great deal of attention. Hence, a local thinker overweights the probability of winning the salient jackpot, and behaves as if she was risk-seeking. In contrast, an agent typically demands insurance against unlikely, but potentially large losses. Compared to the rather small insurance premium the large loss stands out, its probability is inflated, and a local thinker behaves as if she was risk-averse. In fact, the above line of argumentation holds for different models of stimulus-driven attention, that are, *salience theory of choice under risk*

[1] We have borrowed the notion of *local thinking* from a related model by Gennaioli and Shleifer (2010).

(Bordalo *et al.*, 2012, henceforth: BGS) and *a model of focusing* (Kőszegi and Szeidl, 2013, henceforth: KS). Importantly, in the choice contexts described above our modelling of economic salience meets the intuitive notion of salience; that is, whatever is salient in the sense of our model is apparent and not subtle, such as the large potential gains of lottery gambles, or the large outstanding losses an agent will want to insure against. Altogether, models of stimulus-driven attention can account for both a preference for right-skewed and an aversion toward left-skewed risks.

Our contributions in this paper are threefold. First, we show that continuous models of stimulus-driven attention satisfy basic axioms of choice under risk. In particular, for any lottery with finitely many outcomes, there exists a well-defined certainty equivalent that is monotonic in outcomes and probabilities. Kontek (2016) has shown that in discrete model variants certainty equivalents may not exist, and monotonicity in probabilities may be violated. These results hinge on the assumption that in the discrete salience model, for instance, the objective probability of the i^{th} most salient outcome is discounted via a factor δ^{i+1} for some salience-parameter $\delta < 1$. Then, monotonicity in probabilities may be violated if the probability mass is shifted from a low, salient outcome to a larger but less salient outcome which is strongly discounted. BGS use the discrete version of their model for analytical ease. This simplified model is arguably best thought of as an approximation to the more realistic, but also more complex, continuous model. We show that all problems raised by Kontek are resolved in the continuous salience and focusing models.

Second, we show that models of stimulus-driven attention predict skewness preferences. Using the discrete salience model, Bordalo *et al.* (2013a) have argued why individuals like right-skewed and dislike left-skewed assets, but they have not precisely disentangled a local thinker's preferences for risk and skewness. In contrast, we formally derive skewness preferences from continuous models of stimulus-driven attention; that is, we show that a local thinker is more likely to choose a binary risk if it is ceteris paribus (i.e., for a given expected value and variance) skewed further to the right. In addition, we single out the channel—namely, the *contrast effect*—through which models of local thinking predict skewness preferences. The contrast effect means that, when comparing a risky and a safe option, a risky outcome receives the more attention the more it differs from the safe option's payoff. As the models of salience (BGS) and focusing (KS) share the assumption of contrast effects, both predict skewness preferences.

Third, we show that unrealistic predictions of cumulative prospect theory on the magnitude of skewness preferences (e.g., Rieger and Wang, 2006; Azevedo and Gottlieb, 2012; Ebert and Strack, 2015, 2016) can be resolved in the continuous salience and focusing models. For CPT agents, there always exists a sufficiently skewed, small binary risk with negative expected value that is attractive. As a consequence, a CPT agent either gambles until bankruptcy or, if she anticipates her future behavior, does not even gamble when expected gains are arbitrarily large (Ebert and Strack, 2015, 2016). In addition, firms can earn arbitrarily large expected profits by selling skewed lotteries to CPT agents (Rieger and Wang, 2006; Azevedo and Gottlieb, 2012). Notably, discrete versions of the salience and focusing models would also yield the unrealistic predictions on gambling behavior delineated by Ebert and Strack. In contrast, continuous models of stimulus-driven attention predict more plausible behavior.

Skewness preferences are not only relevant for insurance and gambling decisions, but have also important implications for many other economic and financial decision situations. Barberis (2013), for instance, argues that skewness preferences can account for the puzzle that the average return of stocks conducting an initial public offering (IPO) is below that of comparable stocks that did not conduct an IPO. This could be explained by the fact that stocks that conduct an IPO are typically right-skewed and therefore overpriced (Boyer *et al.*, 2010; Bali *et al.*, 2011; Conrad *et al.*, 2013). In this line, Green and Hwang (2012) find that the more skewed the distribution of expected returns is, the lower the long-term average return of an IPO-stock is. Chen *et al.* (2001) even argue that managers strategically disclose information in order to create positive skewness in the distribution of stock returns. This also relates to the well-known growth puzzle (Fama and French, 1992) according to which value stocks, which are underpriced relative to financial indicators, yield higher average returns than (overpriced) growth stocks. Bordalo *et al.* (2013a) suggest that this discrepancy arises as value stocks are typically left-skewed while growth stocks are usually right-skewed. Relatedly, skewness preferences play an important role for portfolio selection (Chunhachinda *et al.*, 1997; Prakash *et al.*, 2003; Mitton and Vorkink, 2007), and allow us to understand the prevalent use of technical analysis for asset trades, even though it is futile in light of the efficient market hypothesis (Ebert and Hilpert, 2016). Finally, a preference for skewness also matters in labor economics. Hartog and Vijverberg (2007) and Berkhout *et al.* (2010) argue that workers accept a lower expected wage if the distribution of wages in a cluster (i.e., education-occupation combination) is right-skewed. In line with this evidence, Choi *et al.* (2016) observe that the number of college students choosing to major in a certain field is the higher the more right-skewed the distribution of stock returns of potential employers is. Altogether, skewness preferences help us to understand various puzzles of economic decision-making.

We proceed as follows. Throughout the paper, we restrict our analysis to the model of salience (BGS) while we establish the analogous results for the focusing model (KS) in Appendix B. In Section 2, we present the continuous salience model. Subsequently, we prove that in this model each discrete lottery has a well-defined certainty equivalent that satisfies monotonicity (Section 3). In Section 4, we show that the salience model predicts skewness preferences. In Section 5, we delineate that puzzles on the magnitude of skewness preferences emerging for CPT agents can be resolved in the salience model. Finally, Section 6 discusses our findings and concludes. All proofs are relegated to Appendix A.

2 Model

According to salience theory of choice under risk, a choice problem is defined by some choice set C, which contains a finite number of lotteries yielding risky monetary payoffs, and the corresponding space of states of the world S. Suppose an agent chooses a lottery from the set $C := \{L_x, L_y\}$ where $L_x := (x_1, p_1; \ldots; x_n, p_n)$ and $L_y := (y_1, q_1; \ldots; y_m, q_m)$ with $n, m \in \mathbb{N}$ and $\sum_{i=1}^{n} p_i = \sum_{i=1}^{m} q_i = 1$. The payoffs x_i denote pairwisely distinct monetary outcomes, which occur with a strictly positive probability $p_i > 0$ for any $1 \leq i \leq n$. If lottery L_x is degenerate (i.e., $n = 1$), we call it a *safe* option. We impose analogous conventions for the outcomes of lottery L_y. Each state of the world $s_{ij} := (x_i, y_j)$ corresponds to a payoff-combination of the available lotteries, and occurs with probability $\pi_{ij} > 0$. A decision-maker evaluates monetary outcomes via a strictly increasing value function $u(\cdot)$ with $u(0) := 0$. We say that the curvature of $u(\cdot)$ reflects a decision-maker's *intrinsic risk-attitude*; that is, she is intrinsically risk-averse (risk-seeking) if $u(\cdot)$ is concave (convex). According to EUT, the expected utility $U(\cdot)$ assigned to lottery L_x equals

$$U(L_x) = \sum_{s_{ij} \in S} \pi_{ij} u(x_i).$$

According to salience theory of choice under risk, a decision-maker evaluates a lottery by assigning a subjective probability to each state s_{ij} that depends on π_{ij} and on the state's salience. In particular, the salience of state $s_{ij} \in S$ is determined by a *salience function* $\sigma(\cdot, \cdot)$, which is symmetric, bounded, continuously differentiable, and satisfies the following three properties:

1. *Ordering*. Let $\mu = \operatorname{sgn}(u(x_i) - u(y_j))$. Then for any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$,

$$\sigma(u(x_i) + \mu\,\epsilon, u(y_j) - \mu\,\epsilon') > \sigma(u(x_i), u(y_j)).$$

2. *Diminishing sensitivity*. Let $u(x_i), u(y_j) \geq 0$. Then for any $\epsilon > 0$,

$$\sigma(u(x_i) + \epsilon, u(y_j) + \epsilon) < \sigma(u(x_i), u(y_j)).$$

3. *Reflection*. For any $u(x_i), u(y_j), u(x_k), u(y_l) \geq 0$, we have

$$\begin{aligned} & \sigma(u(x_i), u(y_j)) < \sigma(u(x_k), u(y_l)) \\ \text{if and only if} \qquad & \sigma(-u(x_i), -u(y_j)) < \sigma(-u(x_k), -u(y_l)). \end{aligned}$$

We say that a state s_{ij} is the more salient the larger its salience value $\sigma(u(x_i), u(y_j))$ is. Thus, the ordering property implies that a state is the more salient the more the lotteries' payoffs in this state differ.[2] In this sense ordering captures the *contrast effect*, according to which a large difference in outcomes within a given state attracts a decision-maker's attention.[3] Diminishing sensitivity reflects *Weber's law* of perception and implies that the

[2]In Appendix A.1 we provide a novel, equivalent definition of the ordering property that is based on the partial derivatives of the salience function (see Lemma 2).

[3]If we fix one argument of the salience function, then the ordering property is equivalent to the contrast effect; that is, the salience of a state increases if and only if the difference in values in this state increases.

salience of a state decreases if the outcomes' values uniformly increase in absolute terms. Hence diminishing sensitivity captures the *level effect* according to which a given contrast in the value of outcomes is more salient for lower outcome levels. Instead of the terms ordering and diminishing sensitivity, we will mainly use the more intuitive notions of contrast and level effects. Throughout the paper, we use $\sigma_{\beta,\theta}(x,y) := \frac{\beta(x-y)^2}{(|x|+|y|+\theta)^2}$ for some $\beta, \theta > 0$ as our leading example of a parametric salience function.

Following the smooth salience characterization proposed in Bordalo *et al.* (2012, page 1255), each state s_{ij} receives the salience weight $\Delta^{-\sigma(u(x_i),u(y_j))}$ for some salience function $\sigma(\cdot,\cdot)$ and some constant $\Delta \in (0,1]$ that captures an agent's susceptibility to salience. A rational decision-maker is captured by $\Delta = 1$, while the smaller Δ is, the stronger the salience bias is. We call an agent with $\Delta < 1$ a *salient thinker*.

Definition 1. *A salient thinker's decision utility $U^s(\cdot)$ for $L_x \in \{L_x, L_y\}$ is given by*

$$U^s(L_x) = \sum_{s_{ij}\in S} \pi_{ij}\, u(x_i) \cdot \frac{\Delta^{-\sigma(u(x_i),u(y_j))}}{\sum_{s_{ij}\in S} \pi_{ij}\, \Delta^{-\sigma(u(x_i),u(y_j))}}.$$

This gives the decision utility according to the continuous model proposed by BGS, where the normalization factor in the denominator ensures that the distorted probabilities sum up to one. Note that for safe options $c \in \mathbb{R}$, we have $U^s(c) = U(c) = u(c)$. Hence, the normalization ensures that a salient thinker's valuation for a safe option c is undistorted, irrespective of the composition of the choice set.

Importantly, the results that we derive in this paper (except for Prediction 2) do not hinge on the assumptions specific to the preceding salience model, but hold for the broader class of models that exhibit contrast effects. We can relax, for instance, the assumption that agents evaluate lotteries based on the objective state space. Indeed, our results would be identical if the salient thinker considers a subset of the state space as long as each outcome of each option is included in (at least) one of this subset's states (see, for instance, the model variant proposed in Dertwinkel-Kalt and Köster, 2015). This is due to the fact that our analysis builds only on choices between a lottery and a safe option. In Appendix B, we further present the analogous results for the closely related focusing model (KS). According to focusing, an agent's attention directed to a given state is determined through a focusing function (i.e., the pendant to the salience function) that satisfies the contrast, but not the level effect. Only Prediction 2 does not rely on the contrast effect, but instead builds on diminishing sensitivity, and is therefore specific to the salience model.

3 Certainty Equivalents and Monotonicity

Models of choice under risk should specify a unique certainty equivalent for any lottery in order to ensure that a lottery's evaluation is well-defined. Certainty equivalents are typically required to satisfy the axiom of monotonicity according to which a lottery's certainty equivalent increases if either probability mass is shifted toward more favorable outcomes or if some outcomes increase. We precisely define these properties as follows.

Definition 2. *Let $L := (x_1, p_1; \ldots; x_n, p_n)$ denote some lottery with $x_i \in \mathbb{R}$ for all $1 \leq i \leq n$. Outcomes are ordered such that $x_1 < \ldots < x_n$, and probabilities $p_1, \ldots, p_n$ sum up to one.*

(a) *The* certainty equivalent *is defined as the minimum monetary sum c that makes a salient thinker indifferent between taking lottery L and getting c for sure. Formally, suppose an agent faces some choice set $\{L, c\}$ comprising a lottery L and a safe option c. Then c is the certainty equivalent to lottery L if and only if $U^s(L) = U^s(c)$.*

(b) *Denote $L' := (x_1, p'_1; \ldots; x_n, p'_n)$ where $p'_i = p_i + \epsilon$ and $p'_l = p_l - \epsilon$ for some $i > l$ and some $0 < \epsilon \leq p_l$ and $p'_k = p_k$ for all $k \neq i, l$. Suppose that c denotes the certainty equivalent to L and c' denotes the certainty equivalent to L'. The certainty equivalent is* monotonic in probabilities *if and only if $c' > c$.*

(c) *Denote $L'' := (x''_1, p_1; \ldots; x''_n, p_n)$ where $x''_l > x_l$ for some $l \in \{1, \ldots, n\}$ and $x''_k = x_k$ for all $k \neq l$. Suppose that c denotes the certainty equivalent to L and c'' denotes the certainty equivalent to L''. The certainty equivalent is* monotonic in outcomes *if and only if $c'' > c$.*

Kontek (2016) establishes that in the discrete salience model certainty equivalents do not satisfy monotonicity in probabilities and may not even exist. We will show that these observations are artefacts of the simplified, discrete salience model that Kontek analyzes. Here, the objective probability of the i^{th} most salient state is discounted via a factor δ^{i+1} for some salience-parameter $\delta < 1$. Therefore, a change in the salience ranking of states induces a discontinuous jump in a salient thinker's valuation for a given lottery. As a consequence, for some lotteries a certainty equivalent may not exist. In addition, for lotteries with more than two outcomes, monotonicity in probabilities may be violated if probability mass is shifted from a low, salient outcome to a larger, but less salient outcome that is strongly discounted.

In order to illustrate why in the discrete salience model for some lotteries a certainty equivalent does not exist, consider the binary lottery that pays \$1 with probability p and \$0 with probability $1-p$. If the lottery's upside of winning \$1 is unlikely (i.e., p is small), a certainty equivalent—being close to the lottery's downside of winning \$0—exists. Here, the lottery's upside is salient. If p increases gradually, the certainty equivalent increases likewise, which implies that the lottery's upside becomes less and its downside becomes more salient. Note, however, that this does not alter the salience weights as long as the salience ranking remains the same (i.e., the upside is still more salient than the downside). There also is some probability $\hat{p}$ for which a certainty equivalent exists and the lottery's up- and downside are equally salient. According to the discrete salience model, a salient thinker's valuation for the above lottery drops discontinuously at $p = \hat{p}$ because for smaller p the lottery's upside is salient, and its probability is overweighted. Hence, there exists some $\epsilon > 0$ such that for any $p \in [\hat{p} - \epsilon, \hat{p})$ no certainty equivalent exists (for a formal analysis, see Kontek, 2016).

BGS apply the simplified, discrete version of their model for analytical ease when the continuous model could be expected to yield identical predictions. In contrast, the above counterintuitive properties of the certainty equivalent rely on the use of the discrete model. We resolve the issues of non-existing and non-monotonic certainty equivalents by investigating the more involved continuous salience model proposed in the previous section. First, we show that given continuous salience distortions each binary lottery has a unique certainty equivalent, which also satisfies monotonicity in probabilities and outcomes. Second, we generalize our findings toward lotteries with finitely many outcomes.

Binary lotteries. Suppose an agent faces a choice set $\{L, c\}$ where $L := (x_1, p; x_2, 1-p)$ is a binary lottery with $x_2 > x_1$ and c denotes the option that pays an amount of $c \in \mathbb{R}$ with certainty. Then, lottery L is (weakly) preferred over the safe option c if and only if

$$U^s(c) \leq U^s(L) = \frac{u(x_1)\, p\, \Delta^{-\sigma(u(x_1),u(c))} + u(x_2)\, (1-p)\, \Delta^{-\sigma(u(x_2),u(c))}}{p\, \Delta^{-\sigma(u(x_1),u(c))} + (1-p)\, \Delta^{-\sigma(u(x_2),u(c))}} =: f(c),$$

while the safe option c is a salient thinker's certainty equivalent to lottery L if and only if

$$c = u^{-1}\left(f(c)\right).$$

For $p = 0$ the certainty equivalent is given by $c = u^{-1}(u(x_2)) = x_2$ while for $p = 1$ it is equal to $c = u^{-1}(u(x_1)) = x_1$. We conclude that the certainty equivalent—given it exists—lies between x_1 and x_2 for any $p \in (0, 1)$ because $u^{-1}(\cdot)$ is strictly increasing and $U^s(L)$ is a convex combination of $u(x_1)$ and $u(x_2)$. Then,

$$u^{-1} \circ f : [x_1, x_2] \to [x_1, x_2], \quad c \mapsto u^{-1}(f(c))$$

is a well-defined continuous function on a closed, convex set which has—by Brouwer's fixed-point theorem—a fixed point. By the ordering property, $\sigma(u(x_1), u(c))$ strictly increases in c, while $\sigma(u(x_2), u(c))$ strictly decreases in c. It follows that $f(c)$ strictly decreases in c, so that the certainty equivalent is unique. Thus, for any $p \in [0, 1]$ a well-defined certainty equivalent c exists.

In order to verify monotonicity in probabilities and outcomes, we define

$$h(x_1, x_2, p, c) := u^{-1}(f(c)) - c$$

where $c = c(x_1, x_2, p)$ denotes the unique certainty equivalent to lottery L. As ordering implies that $\sigma(u(x_1), u(c))$ strictly decreases in x_1 and $\sigma(u(x_2), u(c))$ strictly increases in x_2, we obtain that $f(c)$ strictly increases in x_k for $k \in \{1, 2\}$. Remembering that $f(c)$ strictly decreases in c, we have

$$\frac{\partial h(x_1, x_2, p, c)}{\partial c} < 0 \quad \text{and} \quad \frac{\partial h(x_1, x_2, p, c)}{\partial x_k} > 0, \quad k \in \{1, 2\}.$$

In addition, straightforward computations show that

$$\frac{\partial h(x_1,x_2,p,c)}{\partial p} = \underbrace{u'\left(f(c)\right)^{-1}}_{>0} \cdot \underbrace{\left(-\frac{\Delta_1\Delta_2(u(x_2)-u(x_1))}{(p\Delta_1+(1-p)\Delta_2)^2}\right)}_{<0} < 0,$$

where $\Delta_k := \Delta^{-\sigma(u(x_k),u(c))}$ for $k \in \{1,2\}$. The implicit function theorem then yields

$$\frac{\partial c}{\partial p} = -\frac{\frac{\partial}{\partial p}h(x_1,x_2,p,c)}{\frac{\partial}{\partial c}h(x_1,x_2,p,c)} < 0 \quad \text{and} \quad \frac{\partial c}{\partial x_k} = -\frac{\frac{\partial}{\partial x_k}h(x_1,x_2,p,c)}{\frac{\partial}{\partial c}h(x_1,x_2,p,c)} > 0, \quad k \in \{1,2\}.$$

Hence a salient thinker's certainty equivalent to any binary lottery is well-defined and monotonic in probabilities and outcomes.

Lotteries with finitely many outcomes. We extend our preceding analysis and show that also for a general, discrete lottery $L := (x_1,p_1;\ldots;x_n,p_n)$ with $n \geq 2$ pairwisely distinct outcomes, a certainty equivalent exists and is well-defined. Consider again some choice set $\{L,c\}$, where option c gives the monetary outcome c with certainty. A salient thinker (weakly) prefers lottery L to the safe option c if and only if

$$U^s(c) \leq U^s(L) = \frac{\sum_{i=1}^n p_i\, u(x_i)\Delta^{-\sigma(u(x_i),u(c))}}{\sum_{i=1}^n p_i\, \Delta^{-\sigma(u(x_i),u(c))}} =: f(c).$$

Without loss of generality, we label outcomes such that $x_1 < \ldots < x_n$. A salient thinker's certainty equivalent to lottery L is implicitly given by $c = u^{-1}(f(c))$. By the same arguments as in the case of a binary lottery, the continuous function $u^{-1} \circ f : [x_1,x_n] \to [x_1,x_n]$ has at least one fixed point due to Brouwer's fixed-point theorem, and we obtain the following proposition.

Proposition 1 (Certainty equivalent to a discrete lottery)**.** *A salient thinker's certainty equivalent to a lottery with $n \geq 2$ outcomes is unique and monotonic in outcomes and probabilities.*

For a given lottery L, we can now define a salient thinker's *risk premium* r as the difference in the lottery's expected value $\mathbb{E}[L]$ and its certainty equivalent c, that is $r := \mathbb{E}[L] - c$. Given Proposition 1, a salient thinker's risk premium for lottery L is well-defined. In the next section, we will investigate a salient thinker's risk preferences by determining the size and the sign of her risk premium.

4 Risk Attitudes and Skewness Preferences

In this section, we investigate how salience distortions shape risk attitudes by analyzing under which conditions a salient thinker prefers a lottery over a safe option that pays the lottery's expected value. In Section 4.1, we show that salient thinkers are risk-averse with respect to sufficiently left-skewed lotteries and risk-seeking with respect to sufficiently right-skewed lotteries. This observation can explain, for instance, the simultaneous demand for insurance and casino gambling. We thereby extend findings by BGS (see their Section IV) to the continuous salience model. In Section 4.2, we precisely show that salient thinkers exhibit a preference for skewness. Importantly, we restrict our analysis to binary lotteries since these are uniquely characterized by their first three standardized central moments: expected value, variance, and skewness. While for general lotteries different notions of skewness exist, only for binary gambles skewness is unambigously defined (see Ebert, 2015). Thus, using binary risks, we can precisely analyze a salient thinker's preference over the skewness of lotteries. We relate our findings to the experimental literature on skewness preferences, and derive further testable predictions.

4.1 Stylized Facts on Skewness Preferences

Suppose a decision-maker decides whether to buy some binary lottery L at its fair price. Formally, the decision-maker faces the choice set $\{L, \mathbb{E}[L]\}$ where $L := (x_1, p; x_2, 1-p)$ is a binary lottery with outcomes $x_2 > x_1$, and an expected value $\mathbb{E}[L] := p \cdot x_1 + (1-p) \cdot x_2$. We refer to $\mathbb{E}[L]$ as the actuarially *fair price* of lottery L. In order to deal with indifference, we say that the decision-maker buys the lottery at its fair price if and only if she strictly prefers the risky option L over the safe option $\mathbb{E}[L]$.

In line with BGS, we assume in this section a linear value function $u(x) = x$.[4] Under this assumption, a salient thinker chooses the safe option over the lottery if and only if

$$p \cdot x_1 + (1-p) \cdot x_2 \geq \frac{p \cdot x_1 \cdot \Delta^{-\sigma(x_1, \mathbb{E}[L])} + (1-p) \cdot x_2 \cdot \Delta^{-\sigma(x_2, \mathbb{E}[L])}}{p \cdot \Delta^{-\sigma(x_1, \mathbb{E}[L])} + (1-p) \cdot \Delta^{-\sigma(x_2, \mathbb{E}[L])}}.$$

Rearranging this inequality gives $\Delta^{-\sigma(x_1, \mathbb{E}[L])} \geq \Delta^{-\sigma(x_2, \mathbb{E}[L])}$, or, equivalently,

$$\sigma(x_1, \mathbb{E}[L]) \geq \sigma(x_2, \mathbb{E}[L]).$$

Thus, whenever the lottery's downside x_1 is weakly more salient than its upside x_2, the agent behaves as if she was risk-averse, and chooses the safe option; otherwise, the agent opts for the risky lottery. This highlights a crucial difference in probability weighting under salience and cumulative prospect theory. While the CPT agent overweights small probabilities independent of the corresponding outcome's size, the salient thinker inflates decision weights on salient outcomes.

On the one hand, salience distortions can induce risk-averse behavior. For illustrative reasons, let $x_1 \geq 0$ and $p \leq 1/2$. This immediately implies $\mathbb{E}[L] - x_1 \geq x_2 - \mathbb{E}[L]$; that is, the contrast in the downside payoff and expected value exceeds the contrast in the upside

[4] In contrast to EUT, salience theory does not have to assume a curved value function in order to generate risk-averse or risk-seeking behavior. As salience distortions suffice to generate different risk attitudes, the use of a linear value function is justified (Bordalo *et al.*, 2012).

payoff and expected value. Thus, we obtain

$$\begin{aligned}\sigma(x_1, \mathbb{E}[L]) &> \sigma(\mathbb{E}[L], \mathbb{E}[L] + \mathbb{E}[L] - x_1) \\ &\geq \sigma(\mathbb{E}[L], \mathbb{E}[L] + x_2 - \mathbb{E}[L]) \\ &= \sigma(x_2, \mathbb{E}[L]),\end{aligned}$$

where the first inequality follows from diminishing sensitivity, the second one from ordering, and the final equality from symmetry. We conclude that a salient thinker behaves risk-averse if a non-negative downside payoff is (weakly) less likely than the upside payoff.

On the other hand, a salient thinker sometimes behaves as if she was risk-seeking. As before, suppose $x_1 \geq 0$. If the lottery's upside is unlikely but large compared to its expected value, the salient thinker might buy the lottery at its fair price. In fact, we can construct a binary lottery with a salient upside so that the salient thinker goes for the risky instead of the safe option. The ordering property implies

$$\lim_{p \to 1} \sigma(x_2, \mathbb{E}[L]) = \sigma(x_2, x_1) > \sigma(x_1, x_1) = \lim_{p \to 1} \sigma(x_1, \mathbb{E}[L]).$$

Hence, since the salience function is continuous, there exists some $\hat{p} = \hat{p}(x_1, x_2) \in (1/2, 1)$ such that for any $p > \hat{p}$ the lottery's upside is salient, and the salient thinker chooses the risky option. Due to diminishing sensitivity, a salient thinker behaves as if she was risk-seeking only if the lottery's upside occurs with a strictly lower probability than its non-negative downside. More generally, we obtain the following proposition.

Proposition 2 (Risk attitudes)**.** *Suppose a salient thinker chooses between the binary lottery $L := (x_1, p; x_2, 1-p)$ and the safe option that pays the lottery's expected value. Then, there exists some value $\hat{p} = \hat{p}(x_1, x_2) \in (0, 1)$ such that she chooses the safe option if and only if $p \leq \hat{p}$.*

A straightforward implication of the preceding proposition is that the salience approach accounts for the *fourfold pattern of risk attitudes* (Tversky and Kahneman, 1992). Specifically, a bunch of experimental evidence suggests that people are typically risk-seeking (risk-averse) over low-probability gains (losses), and risk-averse (risk-seeking) over high-probability gains (losses).[5]

Corollary 1 (Fourfold-pattern of risk attitudes)**.**

(a) If $x_2 > x_1 \geq 0$, then $\hat{p} > \frac{1}{2}$.

(b) If $0 \geq x_2 > x_1$, then $\hat{p} < \frac{1}{2}$.

Next, we relate a salient thinker's risk attitude to a lottery's skewness. Ebert (2015) defines the skewness of a binary lottery as its third, standardized central moment

$$S(L) := \mathbb{E}\left[\left(\frac{L - \mathbb{E}[L]}{\sqrt{Var(L)}}\right)^3\right] = \frac{2p - 1}{\sqrt{p(1-p)}} \tag{1}$$

[5]BGS derive a similar result for the discrete salience model.

where $Var(L) := p(1-p)(x_2 - x_1)^2$ denotes the variance of lottery L. Other notions of skewness refer to "long and lean" tails of the risk's probability distribution. Indeed there exist several measures of skewness, which are, however, all equivalent for binary risks (Ebert, 2015, Proposition 2). Thus, only for binary lotteries the impact of skewness on a decision-maker's risk attitude can be unambiguously assessed. In the following, we adopt the short, intuitive notion of skewness which refers to the probability that the lottery's downside payoff is realized.

Definition 3 (Skewness of binary risks)**.** *Consider two binary lotteries $L_x := (x_1, p; x_2, 1-p)$ and $L_y := (y_1, q; y_2, 1-q)$ with $x_2 > x_1$ and $y_2 > y_1$. We say that L_x is more (less, equally) skewed than L_y if and only if $p > q$ ($p < q, p = q$). Lottery L_x is called right-skewed if $p > \frac{1}{2}$, left-skewed if $p < \frac{1}{2}$, and symmetric otherwise.*

From Equation (1) it is straightforward to see that $S < 0$ for any left-skewed lottery, $S > 0$ for any right-skewed lottery, and $S = 0$ for any symmetric lottery. Therefore, we also say that a left-skewed (right-skewed) lottery is *negatively* (*positively*) skewed, and that a lottery is the more skewed the larger S is.

The distribution of various downside risks such as car accidents or natural disasters is typically left-skewed: these events are rare, but if they happen they are severe. In this context, option $\mathbb{E}[L]$ may reflect a fair-priced insurance contract against the downside risk. The distribution of casino gambling, or lottery games, on the other hand, is typically right-skewed: gains are large, but occur rarely. Here, option $\mathbb{E}[L]$ can be interpreted as the fair price to bet on an upside risk.

The finding that agents seek right-skewed risks but tend to avoid left-skewed risks is established in the literature as skewness preferences. A tendency to choose right-skewed risks has been observed by Golec and Tamarkin (1998) with respect to horse-race betting, by Garrett and Sobel (1999) in the context of lottery games, and in several studies on investment behavior (Boyer *et al.*, 2010; Bali *et al.*, 2011; Green and Hwang, 2012; Conrad *et al.*, 2013). At the same time, consumers insure against left-skewed risks as demonstrated by Sydnor (2010) and Barseghyan *et al.* (2013) who analyze deductible choices in auto and home insurance contracts. The following stylized examples illustrate that salience theory can account for this empirical evidence.[6]

Example 1 (Insurance). Suppose the agent has to decide whether to pay the fair insurance premium in order to avoid a binary risk L. In a typical insurance example, the risky option yields a large loss (i.e., $x_1 < 0$) with a small probability and zero payoff (i.e., $x_2 = 0$) otherwise. Then, according to Proposition 2, a salient thinker buys the insurance if the probability of the loss is sufficiently small.

[6] Notably, salience theory can also explain the demand for small scale insurance, e.g. insurance for consumption goods such as TVs or smartphones, where the potential loss is high relative to the insurance premium but not large overall. Cicchetti and Dubin (1994), for instance, report that many consumers pay a substantial premium in order to avoid the small risk (less than one percent) of having to pay $55 for repair in case their internal telephone wiring breaks down.

Example 2 (Gambling). Suppose the agent decides whether to buy a lottery ticket at its fair price. When participating in the lottery, she could win either a large amount (i.e., $x_2 > 0$) or nothing (i.e., $x_1 = 0$). The salient thinker might prefer the gamble, but due to diminishing sensitivity only if the risk is right-skewed. According to Proposition 2, the salient thinker buys the lottery ticket if the probability of the gain is sufficiently small.

Example 3 (Investments). Suppose the agent decides whether to buy an asset—that either pays $x_1 < 0$ or $x_2 > 0$ in the future—at its fair price. If the probability of the gain is sufficiently high, the downside payoff x_1 stands out and the salient thinker does not invest in the asset. If the probability of the loss is high, the upside payoff x_2 is salient, and the decision-maker buys the asset at its fair price. As Bordalo *et al.* (2013a) have already pointed out, this implies a tendency to buy right-skewed assets.[7]

4.2 Salience and Skewness Preferences

In line with the presented empirical evidence, salience theory suggests that the skewness of a risk's probability distribution affects an agent's attitude toward risk. Most field studies, however, do not precisely test for the role of skewness in risk-taking since the variance and skewness of typical casino gambles or lottery games are not independent, but are highly correlated. Thus, risk and skewness preferences cannot be disentangled. Ebert (2015) argues, for instance, that inferring skewness preferences at the horse track from the study by Golec and Tamarkin (1998) might be misleading. In fact, increasing the skewness of a stylized horse race bet $L = (1/p, p; 0, 1-p)$, while holding its expected value and variance (i.e., the corresponding risk) constant, does not yield a new horse race bet, but a lottery with very different properties. Ebert (2015) concludes that "a choice between two horse-race bets is never a choice between different levels of skewness only." Indeed, for any given outcomes x_1 and x_2 a change in the lottery's probability distribution induces a change in its expected value $\mathbb{E}[L]$ and variance $Var(L)$. As a consequence, we cannot infer from Proposition 2 whether it is solely the skewness of a risk that induces a salient thinker's aversion toward left-skewed and her preference for right-skewed lotteries. In order to disentangle a salient thinker's preference for skewness from her preference for risk, a lottery's skewness needs to be varied for a fixed expected value and variance.

Lemma 1 (Ebert's moment characterization of binary risks)**.** *For constants $E \in \mathbb{R}$, $V \in \mathbb{R}_+$ and $S \in \mathbb{R}$, there exists exactly one binary lottery $L = (x_1, p; x_2, 1-p)$ with $x_2 > x_1$ such that $\mathbb{E}[L] = E$, $Var(L) = V$ and $S(L) = S$. Its parameters are given by*

$$x_1 = E - \sqrt{\frac{V(1-p)}{p}}, \; x_2 = E + \sqrt{\frac{Vp}{1-p}}, \; \textit{and } p = \frac{1}{2} + \frac{S}{2\sqrt{4+S^2}}. \tag{2}$$

[7]While Bordalo *et al.* (2013a) state that salience predicts a "taste for skewness" in the context of asset choices, we will precisely disentangle a salient thinker's preferences for risk and skewness. Thereby, we are the first to formally derive a salient thinker's preference for skewness.

For a proof of Lemma 1 see Ebert (2015). In the following, we will refer to the unique binary lottery that has expected value E, variance V, and skewness S as $L(E, V, S)$. Using the above moment characterization of binary risks, we can assess the impact of skewness on a salient thinker's risk attitude. As before, we assume a linear value function $u(x) = x$ so that the salient thinker's risk premium for the binary lottery $L(E, V, S)$ equals

$$r(E, V, S) = \sqrt{Vp(1-p)} \cdot \left(\frac{\Delta^{-\sigma(x_1,E)} - \Delta^{-\sigma(x_2,E)}}{p\Delta^{-\sigma(x_1,E)} + (1-p)\Delta^{-\sigma(x_2,E)}} \right), \tag{3}$$

where outcomes $x_k = x_k(E, V, S)$, $k \in \{1, 2\}$, and probability $p = p(S)$ are defined in Equation (2). A salient thinker strictly prefers the risky option $L(E, V, S)$ over the safe option E if and only if the lottery's risk premium is strictly negative, or, equivalently, its upside payoff is salient.

Proposition 3 (Skewness preferences)**.** *For a given expected value E and variance V, there exists a unique skewness value $\hat{S} = \hat{S}(E, V) \in \mathbb{R}$ such that $r(E, V, \hat{S}) = 0$. A salient thinker strictly prefers the binary lottery $L(E, V, S)$ over its expected value E if and only if $S > \hat{S}$.*

Suppose the lottery's expected value and variance are fixed. Then, by Equation (2), increasing the lottery's skewness S increases the probability that its downside payoff is realized. If the lottery's downside payoff becomes more likely, the difference between its upside payoff and the expected value increases, thereby making the lottery's upside more salient. At the same time, the difference between the downside payoff and the expected value decreases so that the lottery's downside becomes less salient. Hence, a salient thinker is the more likely to take a binary risk the more skewed this risk is.[8] By continuity of the salience function we obtain the following corollary.

Corollary 2. *For a given expected value E and variance V, there exists a sufficiently skewed binary lottery for which a salient thinker is willing to pay more than its fair price E.*

As the salience function is bounded, we further conclude from Equation (3) that the risk premium converges to zero if the lottery's skewness becomes arbitrarily large or small.

Corollary 3. *For any expected value E and variance V, we have $\lim_{S \to \pm\infty} r(E, V, S) = 0$.*

Skewness preferences and the contrast effect. Intuitively, in the salience model, skewness preferences are driven by the contrast effect. The stronger the contrast effect is, the more pronounced is a large difference between a lottery's payoff and its expected value. For a positively skewed lottery, the upside payoff differs by more from the expected value than the downside payoff, while the opposite holds for a negatively skewed lottery. Therefore, if the contrast effect becomes stronger, a salient thinker's preference for positive skewness is enhanced. We formalize this idea as follows.

[8] In principle, also EUT could explain skewness preferences. In order to match evidence on risk-averse behavior, however, EUT needs to assume a concave value function. Under this assumption EUT cannot explain why individuals seek right-skewed but avoid left-skewed risks.

Definition 4. *We say that the contrast effect is stronger for salience function σ than for salience function $\hat{\sigma}$ if for any $y \in \mathbb{R}$ the difference $\sigma(x, y) - \hat{\sigma}(x, y)$ is increasing in $|x - y|$.*

The contrast between two values is typically measured by their difference. In this sense, the preceding definition captures the intuitive notion that the contrast effect is stronger for one salience function than another if their difference (i.e., the difference in salience values) increases in the difference of their arguments.

Proposition 4 (Contrast and skewness preferences)**.** *Let the contrast effect be stronger for salience function σ than for salience function $\hat{\sigma}$. Then, a salient thinker's risk premium $r(E, V, S)$ is larger for σ than for $\hat{\sigma}$ if and only if the lottery is left-skewed.*

This suggests that a stronger contrast effect enhances a salient thinker's aversion toward left-skewed risks and her preference for right-skewed risks. More precisely, the preceding proposition implies that the skewness threshold $\hat{S}$ defined in Proposition 3 lies the closer to zero the stronger the contrast effect is. Since we derive the preference for skewness from lotteries with the same expected value, the salience function's second argument is held fixed so that the contrast effect is equivalent to the ordering property in this context. In other words, a salient thinker's preference for skewness is the stronger the more important ordering is relative to diminishing sensitivity.

Since a model of focusing (KS) also builds on the contrast effect, it shares all of our central results on skewness preferences (see Appendix B for a formal proof). In contrast, the model of relative thinking by Bushong *et al.* (2016), which builds on the setup by KS, but assumes reverse contrast effects (i.e., attention assigned to a state decreases in the range of payoffs in this state), cannot account for skewness preferences.

Experimental evidence on skewness preferences. Our preceding results are in line with experimental evidence on skewness-seeking choices. In contrast to studies with field data, laboratory experiments allow us to precisely test for a subject's preference for positive skewness (i.e., the skewness of a lottery can be varied ceteris paribus). Ebert and Wiesen (2011) find that a majority of subjects chooses a right-skewed over a left-skewed binary lottery with the same expected value and variance.[9] Ebert (2015) confirms this preference for right-skewed over left-skewed binary risks. In addition, he observes that a majority of subjects who have to choose between a symmetric and a right-skewed lottery, which has the same expected value and variance, opt for the more skewed alternative. If the choice is between a symmetric and a left-skewed lottery, subjects tend to avoid the left-skewed risk, thereby again choosing the more skewed lottery. Further studies using binary (e.g.,

[9]More precisely, subjects have to choose between two binary lotteries that form a *Mao pair* (Mao, 1970). For any $p \in (0, 1/2)$, two perfectly correlated, binary lotteries $L_x := (x_1, p; x_2, 1-p)$ and $L_y := (y_1, 1-p; y_2, p)$ form a Mao pair if both have the same expected value and variance. The lotteries of a Mao pair differ only in their skewness (Ebert and Wiesen, 2011). Lottery L_x is left-skewed (i.e., its high payoff x_2 occurs with a high probability), while lottery L_y is right-skewed (i.e., its high payoff y_2 occurs with a small probability). In line with Definition 3, Ebert and Wiesen (2011) state that "an individual is said to be *skewness seeking* if, for any given Mao pair, she prefers L_y over L_x." In Appendix C we prove that, for any Mao pair, a salient thinker prefers L_y over L_x.

Brünner *et al.*, 2011) or more complex lotteries (e.g., Grossman and Eckel, 2015) report similar results on skewness-seeking choices. In line with Proposition 3, Åstebro *et al.* (2015) observe that subjects tend to make riskier decisions if the choice set includes right-skewed lotteries. Altogether, a substantial body of research documents skewness-seeking choices under controlled conditions in the laboratory.

Notably, none of the existing laboratory studies has explicitly tested our main prediction that arises from Proposition 3. According to the salience approach, there exists a certain threshold value $\hat{S}$ so that a subject prefers a binary lottery over its expected value if and only if the lottery's skewness exceeds this threshold value. This prediction is novel, and experimentally testable:[10] in detail, fix some expected value E and some variance V. Let subjects choose repeatedly between the safe option paying E and the binary lottery $L(E, V, S)$, where the lottery's skewness S is gradually increased. We hypothesize that the subjects' choices are monotonic in the sense that a subject should opt for the safe option if the lottery's skewness falls below a certain threshold, and for the lottery otherwise.

Prediction 1 (Skewness preferences)**.**

(a) Suppose $S' < S$. If a subject chooses E from the set $\{L(E, V, S), E\}$, she also chooses E from the set $\{L(E, V, S'), E\}$.

(b) Suppose $S'' > S$. If a subject chooses $L(E, V, S)$ from the set $\{L(E, V, S), E\}$, she also chooses $L(E, V, S'')$ from the set $\{L(E, V, S''), E\}$.

While Prediction 1 relates to a within-subjects design, the predictions for a between-subjects experiment are straightforward. As the threshold value defined in Proposition 3 depends on the curvature of the salience function, it will arguably vary across subjects. Then, Prediction 1 implies that the share of subjects choosing the binary lottery $L(E, V, S)$ over its expected value E monotonically increases in the lottery's skewness S.

If we impose more structure on the salience function, we can derive further experimentally testable predictions. In order to predict how subjects' choices depend on the lottery's expected value, for instance, we have to be more precise on how the relative importance of the contrast and level effects vary with the payoff level.

Definition 5 (Decreasing diminishing sensitivity)**.** *A salience function satisfies* decreasing diminishing sensitivity *if and only if, for any values $x, y, z > 0$ such that $x \geq z$ and $\sigma(x - z, x) \geq \sigma(x + y, x)$, the differences $\sigma(x - z, x) - \sigma(x + y, x)$ and $\sigma(-x + z, -x) - \sigma(-x - y, -x)$ are strictly decreasing functions in x.*

A salience function satisfies decreasing diminishing sensitivity if the level effect becomes weaker as the lottery's payoffs increase in absolute terms. Accordingly, diminishing sensitivity is more important at low rather than high payoff levels. As we delineate

[10] BGS as well as Frydman and Mormann (2017) conduct related experiments, but simultaneously vary the lottery's variance and skewness. Hence, these experiments do not test our novel prediction on salience-induced skewness preferences.

in Appendix A.3, a wide class of salience functions—to the best of our knowledge, any salience function that has been proposed in the literature—satisfies decreasing diminishing sensitivity. This property allows us to make the following prediction.

Prediction 2 (Level-dependent skewness preferences)**.**

(a) Suppose $E' < E$ ($E' > E$). If a subject chooses E from the set $\{L(E, V, S), E\}$ and the lottery's payoffs are positive (negative), she also chooses E' from the set $\{L(E', V, S), E'\}$.

(b) Suppose $E'' > E$ ($E'' < E$). If a subject chooses $L(E, V, S)$ from $\{L(E, V, S), E\}$ and the lottery's payoffs are positive (negative), she chooses $L(E'', V, S)$ from $\{L(E'', V, S), E''\}$.

The preceding prediction states that a binary lottery with positive payoffs becomes the more attractive the larger its expected value is. If the expected value and therefore the payoff level increases, diminishing sensitivity becomes weaker, the contrast effect becomes relatively more important, and an agent's preference for right-skewed lotteries is enhanced. Our salience approach therefore predicts that the share of subjects choosing the risky option increases in the lottery's expected value. While we are the first to formally derive this prediction, it has been experimentally confirmed in a recent study by Frydman and Mormann (2016, Table II.A).[11] The prediction for lotteries with negative payoffs is reversed, and has not been tested yet. Formally, Prediction 2 follows from Proposition 7 that is proven in Appendix A.3.

Finally, Corollary 3 yields another testable implication: for any expected value E and variance V, there exists a certain skewness level such that beyond this level the lottery's certainty equivalent decreases in the lottery's skewness. Specifically, the certainty equivalent to $L(E, V, S)$ approaches the lottery's expected value E if the lottery's skewness S becomes arbitrarily large. This gives rise to the following prediction.

Prediction 3 (Limits of skewness preferences)**.** *Consider choice set $\{L(E, V, S), E + \epsilon\}$ for some $\epsilon > 0$. There exists some skewness value $\tilde{S} \in \mathbb{R}$ such that for any $S \geq \tilde{S}$ the subject chooses $E + \epsilon$ over $L(E, V, S)$.*

Importantly, all predictions carry over to the case where only a share of subjects has a linear value function. As we will delineate in the following section, some subjects might have a concave value function, and might therefore never choose a binary lottery (with positive payoffs) over its expected value. In this case all predictions continue to hold: first, those subjects that reveal a linear or close-to-linear value function accord with Prediction 1, while the remaining subjects should always go for the safe option. Second, Prediction 2 should hold as long as the third derivative of the value function is not too negative.[12]

[11] Unfortunately, the authors do not report these results anymore in their new working paper version (Frydman and Mormann, 2017).

[12] Otherwise, if the lottery's expected value increases, its payoffs fall in a range where the value function is much more concave and where the decision-maker is therefore intrinsically much more risk-averse. This effect may countervail the illustrated salience effect so that the lottery may become less attractive due to the increase in its expected value.

Third, Prediction 3 holds for any decision-maker with a (weakly) concave value function. Thus, our predictions are robust to the assumption of heterogeneous agents who differ in the shape of their value functions (i.e., with respect to their degree of intrinsic risk aversion).

The preceding predictions further allow us to test the fundamental assumptions of the salience model. Prediction 1 is driven by the ordering property and the contrast effect (that also the focusing model includes), while Prediction 2 is an implication of diminishing sensitivity and the level effect (that is specific to the salience model). As a consequence, the predictions above are not only valuable in order to test the predictive power of the salience and focusing models with respect to skewness preferences, but also to distinguish between these different approaches to stimulus-driven attention.

5 Puzzles on Skewness Preferences

In many respects, the predictions by salience theory of choice under risk coincide with the predictions by cumulative prospect theory (for a detailed discussion, see BGS). For instance, both theories predict that whether an agent buys insurance or prefers to gamble depends on the skewness of the risk's underlying probability distribution. As shown by three articles, however, cumulative prospect theory yields unreasonably strong predictions on the impact of skewness on risk-taking. On the one hand, Ebert and Strack (2015) argue that for any value function, there exists a right-skewed and arbitrarily small binary risk with a negative expected value that is attractive to a CPT agent. This results in unrealistic predictions for dynamic investment or gambling decisions. On the other hand, Rieger and Wang (2006) as well as Azevedo and Gottlieb (2012) delineate that under "virtually all functional forms that have been proposed in the literature" (Azevedo and Gottlieb, 2012, page 1294) an CPT-agent's willingness to pay for a binary lottery with a fixed expected value is unbounded (i.e., it becomes arbitrarily large if the lottery's upside payoff becomes arbitrarily large). In the following, we will compare salience and cumulative prospect theory's predictions on skewness preferences in the small (Ebert and Strack, 2015) and in the large (Rieger and Wang, 2006; Azevedo and Gottlieb, 2012).

5.1 Skewness Preferences in the Small

Consider a dynamic setup where a decision-maker gambles according to the following strategy: she decides to start gambling, but will stop as soon as she has realized either a rather small loss x_1 or a large gain x_2. This stopping strategy with two absorbing endpoints can be represented as a binary lottery that gives a small loss with a large probability, and a large gain with a small probability. According to Corollary 1, a salient thinker with a linear value function is willing to pay more than the fair price to enter the corresponding gamble if this binary risk is sufficiently skewed. If the decision-maker is naïve and cannot commit to a long-run stopping strategy, but can revise her strategy after every single gain or loss, she never stops gambling as she can always construct a sufficiently skewed stopping strategy that attracts her. Independent of previous gains or losses, a salient thinker decides to gamble in every period anew and therefore continues until bankruptcy.

Likewise, CPT agents that cannot commit to a certain gambling strategy will gamble "until the bitter end" (Ebert and Strack, 2015). Ebert and Strack show that without commitment a naïve CPT agent who uses the preceding stopping strategy will never stop gambling irrespective of her value function's curvature.[13] In particular, Ebert and Strack (2015) verify that CPT agents reveal *skewness preferences in the small*; that is, sufficiently right-skewed binary lotteries with outcomes x_1 and x_2 that are sufficiently small in absolute terms are attractive even if these lotteries' expected values are negative. For these lotteries probability weighting may predominate loss aversion so that the CPT agent participates in an unfair gamble.

While also salient thinkers might gamble until the bitter end, the lotteries which are attractive to a salient thinker are fundamentally different. An attractive lottery's downside payoff should be close to the lottery's expected value, therefore being non-salient. At the same time, the upside payoff should be very large, thereby exceeding the expected value by much in order to stand out and attract the decision-maker's attention. Thus, it is not a preference for skewness in the small that induces a salient thinker to gamble until bankruptcy. It is a preference for lotteries with a large, outstanding upside payoff, which we regard as the more plausible driver of taking up unfair gambles. Forrest *et al.* (2002) precisely capture this intuition by stating that the purchase of a lottery ticket corresponds to "buying a dream." A decision-maker might dream of winning the large jackpot, which allows her to quit her tedious job or to buy an expensive car, thereby overweighting the probability that her dream will come true.

Cumulative prospect theory's prediction that an agent will, irrespective of her value function, gamble until bankruptcy has been regarded as implausible and therefore as a weakness of the model. We will show that this prediction does not necessarily hold for salient thinkers. In fact, it depends on the interplay of a salient thinker's value and salience functions whether she is inclined to gamble or not.[14] Hence, in the salience model agents with a rather linear value function may gamble until the bitter end, while those with a strongly curved value function may not. We regard it as plausible to assume that people are heterogeneous with respect to the curvature of their value functions since some are

[13] The naïve agent does not anticipate that she will not stick to her initial plan in the future. At every point in time, she constructs a new, attractive gambling strategy with a positive skew and a negative expected value, and continues gambling until she has lost her entire wealth. In contrast, a sophisticated agent who cannot commit to future behavior never starts to gamble (Ebert and Strack, 2016) as she is aware of her time-inconsistency. Hence, she foresees that if she adopts the preceding strategy, she will not follow it until the end, but she will stop when her gains come close to the strategy's upper stopping threshold. At this point, following the preceding strategy represents a left-skewed gamble that the CPT agent wants to avoid. Since the agent anticipates that she will stop *too early*, she decides not to gamble in the first place. Importantly, she does not even start to gamble if the expected gains from gambling become arbitrarily large.

[14] The fundamentals of the salience model, that is, the value function u, the salience function σ, and the salience parameter Δ can be estimated simultaneously from real choice data as they are not perfectly collinear. Dertwinkel-Kalt *et al.* (2017a), for instance, conduct such an estimation for the closely related focusing model, simultaneously estimating the value and the focusing function.

intrinsically more risk-averse than others. Given this type of heterogeneity among agents the salience approach predicts that only some people gamble excessively (those with a value function that is linear or close-to-linear) while others do not.[15]

Static salience predictions. Suppose a salient thinker faces some choice set $\{L, \mathbb{E}[L]\}$. For simplicity and in line with the gambling example, let $x_2 > x_1 \geq 0$. We relax our previous assumption of a linear value function, and assume that the decision-maker's value from money is strictly increasing and weakly concave, that is, $u'(\cdot) > 0$ and $u''(\cdot) \leq 0$. As before, we normalize $u(0) := 0$. Then, a salient thinker strictly prefers the risky lottery L over the safe option $\mathbb{E}[L]$ if and only if

$$\frac{u(x_2) - u(\mathbb{E}[L])}{u(\mathbb{E}[L]) - u(x_1)} \cdot \frac{1-p}{p} > \frac{\Delta_1}{\Delta_2},$$

where $\Delta_k := \Delta^{-\sigma(u(x_k), u(\mathbb{E}[L]))}$ for $k \in \{1, 2\}$. For any given expected value $E = \mathbb{E}[L]$, substituting $p = (x_2 - E)/(x_2 - x_1)$ yields

$$\frac{\frac{u(x_2)-u(E)}{x_2-E}}{\frac{u(E)-u(x_1)}{E-x_1}} > \frac{\Delta_1}{\Delta_2}. \tag{C.1}$$

The left-hand side of this inequality constitutes the ratio of the secants' slopes through the points $(E, u(E))$ and $(x_k, u(x_k))$ for $k \in \{1, 2\}$, which is less or equal than one for any weakly concave value function. The right-hand side of Inequality (C.1) gives the ratio of the salience weights which is below one if and only if the lottery's upside is salient. Analogously to the previous section, we can conclude that the lottery's downside is salient whenever the lottery is left-skewed or symmetric.[16] While there exists a right-skewed lottery with a salient upside for any value function, it remains uncertain whether a salient thinker buys this lottery or not.

Intuitively, one would expect that Condition (C.1) is less likely to hold if the value function's curvature increases since intrinsic risk aversion becomes stronger. In addition, compared to a linear value function, the contrast between the values assigned to the lottery's upside payoff and expected value is reduced. As the preference for skewness is driven by the contrast effect, it follows that salience distortions are weaker and therefore less likely to induce risk-seeking behavior if the value function is concave. Indeed, the left-hand side of Condition (C.1) decreases in the value function's curvature; but the corresponding effect on the ratio of salience weights is ambiguous as it depends on how the

[15]Under the assumption that some people are sophisticated while others are naïve, also cumulative prospect theory predicts heterogeneous, but implausible gambling behavior: while naïve agents gamble until bankruptcy (Ebert and Strack, 2015), sophisticated agents do not even gamble if the expected profit becomes arbitrarily large (Ebert and Strack, 2016). For the salience model, in contrast, heterogeneity in gambling behavior follows directly from the model's fundamentals (i.e., the curvature of the value function), and results in more plausible predictions.

[16]Note that $u(E) - u(x_1) \geq u(x_2) - u(E)$ for any $p \leq 1/2$ due to (weak) concavity of the value function. Then, diminishing sensitivity implies that the lottery's downside is weakly more salient than its upside since $u(x_2) > u(x_1) \geq 0$ holds by assumption.

relative importance of ordering and diminishing sensitivity change with the level of values assigned to the outcomes. Therefore, it is impossible to make a general statement on how the value function's curvature affects a salient thinker's risk attitude (see Example 5 for an illustration).

More can be said about the properties of the salience function that facilitate risk-seeking behavior. Under the assumption of a linear value function, as established in Proposition 4, a salient thinker's preference for right-skewed risks is driven by the contrast effect. Intuitively, a salient thinker is especially prone to gamble (even at an unfair price) if a large gain occurring with a small probability stands in sharp contrast to the lottery's expected value, thereby grabbing a great deal of attention. Hence, a salient thinker is the more risk-seeking with respect to sufficiently right-skewed, binary lotteries the stronger the contrast effect is relative to the level effect. In order to verify that this intuition carries over to the case of a concave value function, we compare salience functions that differ in the strength of the contrast effect.

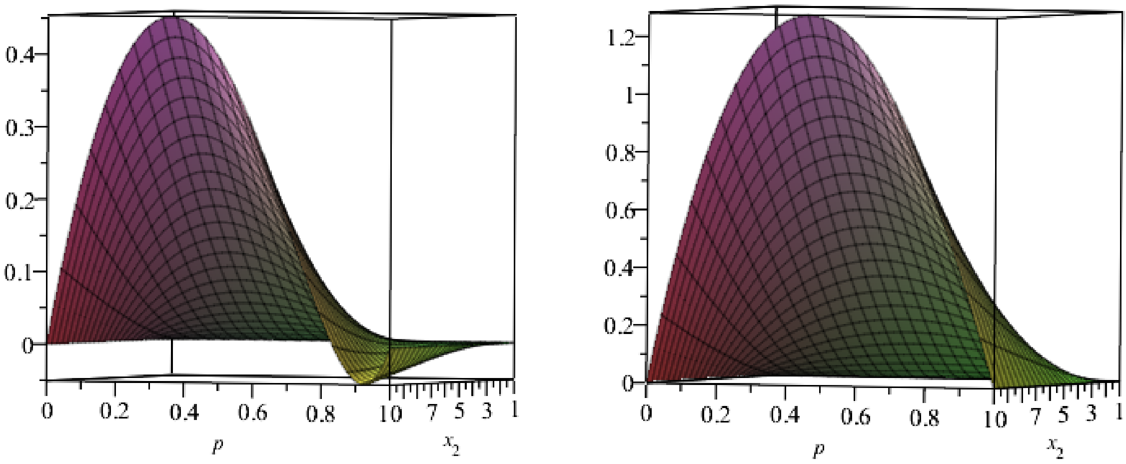
Proposition 5. *Let the contrast effect be stronger for salience function σ than for salience function $\hat{\sigma}$. If lottery L satisfies (C.1) for salience function $\hat{\sigma}$, it also satisfies (C.1) for salience function σ.*

If an agent's value function is very concave and her salience function exhibits a weak contrast effect, we expect that there exists no binary lottery that the agent prefers to its expected value (i.e., Condition (C.1) is never satisfied). We show this with the use of two examples for which we assume power utility $u(x) = x^\alpha$ with $\alpha \in (0, 1)$ and our standard salience function $\sigma_{\beta,\theta}(x, y)$ with $\beta, \theta > 0$. Let $\theta = 0.1$ and $\Delta = 0.7$.

Example 5 (Value function). For a linear value function, the left-hand side of (C.1) equals one, and the salient thinker chooses a lottery whenever its upside is salient. This lottery exists by Proposition 3. Then, due to continuity, Condition (C.1) also holds for a mildly concave value function $u(x) = x^\alpha$ with α being close to one. Let $\beta = 1$ so that the salience function is $\sigma_{\beta,\theta}(x, y) = \frac{(x-y)^2}{(|x|+|y|+0.1)^2}$. If the value function's curvature increases, that is, parameter α decreases, we observe that (C.1) is less likely to hold. More specifically, numerical computations show that there exists some threshold value $\hat{\alpha} \in (0, 1)$ such that for any $\alpha \in (0, \hat{\alpha})$ no unfair, attractive gamble exists. For $\alpha = 0.95$ and $\alpha = 0.5$, Figure 1 depicts the risk premium r as a function of probability p and upside payoff x_2 for $x_1 = 1$.

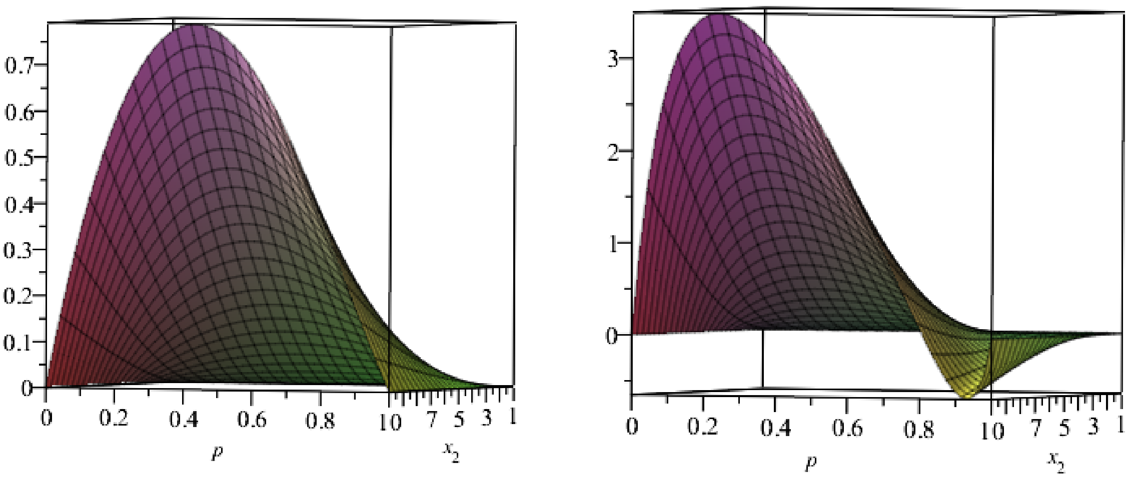
Example 6 (Salience function). Fix $\alpha = 3/4$ so that the value function is $u(x) = x^{3/4}$. We observe that Inequality (C.1) is more likely to hold for at least some binary lottery L if parameter β increases.[17] In fact, numerical computations show that there exists some $\hat{\beta} > 1$ such that for any $\beta > \hat{\beta}$ at least one unfair, attractive gamble exists. For $\beta = 1$ and $\beta = 10$, Figure 2 illustrates the risk premium r as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.

[17] The larger β is the stronger is the contrast effect for salience function $\sigma_{\beta,\theta}$. Indeed, this holds only using the following notion of a stronger contrast effect which is weaker than that stated in Definition 4: for any $\beta > \tilde{\beta}$ and $x, y, z \in \mathbb{R}$, we have $\sigma_{\beta,\theta}(x, z) - \sigma_{\tilde{\beta},\theta}(x, z) > \sigma_{\beta,\theta}(y, z) - \sigma_{\tilde{\beta},\theta}(y, z)$ if $x > y \geq z$ or $x < y \leq z$.



Risk premium for $\alpha = 0.95$. *Risk premium for* $\alpha = 0.5$.

Figure 1: *The above graphs show the risk premium as a function of the upside payoff* x_2 *and the probability* p *that the downside payoff* x_1 *is realized. For* $\alpha = 0.95$*, the risk premium becomes negative for highly right-skewed lotteries (i.e., a large probability* p *on the downside payoff* x_1*) with a large upside payoff* x_2*. For* $\alpha = 0.5$*, the risk premium is non-negative for any feasible lottery.*



Risk premium for $\beta = 1$. *Risk premium for* $\beta = 10$.

Figure 2: *The above graphs show the risk premium as a function of the upside payoff* x_2 *and the probability* p *that the downside payoff* x_1 *is realized. For* $\beta = 10$*, the risk premium becomes negative for highly right-skewed lotteries (i.e., a large probability* p *on the downside payoff* x_1*) with a large upside payoff* x_2*. For* $\beta = 1$*, the risk premium is non-negative for any feasible lottery.*

Comparison to the discrete salience model. For the discrete salience model there always exists an unfair, binary lottery with a salient upside that is attractive to a salient thinker. This result is driven by the fact that for a lottery with a salient upside the right-hand side of Inequality (C.1) simplifies to the salience-parameter $\delta < 1$ (as introduced in the discussion of the discrete salience model after Definition 2). Therefore, the right-hand side of Inequality (C.1) is bounded away from one, while its left-hand side approaches one if the variance of lottery L goes to zero. Thus, the resolution of Ebert and Strack's skewness puzzles relies on the use of the continuous salience model.

5.2 Skewness Preferences in the Large

Rieger and Wang (2006) and Azevedo and Gottlieb (2012) show that cumulative prospect theory also yields implausible predictions on the profitability of selling right-skewed lotteries with large absolute payoffs. Denote $\mathcal{L}(E)$ as the set of all binary lotteries with expected value $E \in \mathbb{R}$. Azevedo and Gottlieb (2012) argue that the expected profit that can be earned by selling a lottery $L \in \mathcal{L}(E)$ to a CPT agent may be unbounded. This prediction arises since the assumption of non-linear probability weighting might induce an unbounded valuation for a lottery with a finite expected value (Rieger and Wang, 2006). If small probabilities are overweighted, increasing a lottery's upside payoff, and reducing the corresponding probability can make this lottery more attractive. This allows a firm to realize an arbitrarily large expected profit if it offers a binary lottery with an arbitrarily large upside payoff (*skewness preferences in the large*).

We show that this puzzle can be resolved for salient thinkers. As before, suppose the decision-maker has a (weakly) concave value function and faces some choice set $\{L, z\}$, where z denotes the price of lottery L. The agent buys the lottery as long as it is strictly preferred over the monetary sum z. Since the salience function is bounded, there exists some threshold value $\bar{\Delta} < \infty$ such that $\Delta^{-\sigma(x,y)} < \bar{\Delta}$ for any $(x, y) \in \mathbb{R}^2$. The following proposition states that for any expected value E, the price a salient thinker is willing to pay for lottery $L \in \mathcal{L}(E)$ is bounded.

Proposition 6. *A salient thinker's valuation for any lottery $L \in \mathcal{L}(E)$ is bounded from above by a function which is affine in E.*

Suppose a firm offers a binary lottery $L \in \mathcal{L}(E)$ at some price z. Optimally, it will set a price equal to the lottery's certainty equivalent, which is well-defined according to Proposition 1. Therefore, the firm will, for a given expected value E, choose to sell the lottery $L \in \mathcal{L}(E)$ that has the largest certainty equivalent. According to Proposition 6, this certainty equivalent is bounded so that the expected profit a firm can earn from selling a binary lottery with a finite expected value cannot become arbitrarily large.[18]

[18] Indeed the expected profit that can be earned from selling a lottery $L \in \bigcup_{E \in \mathbb{R}} \mathcal{L}(E)$ is unbounded.

6 Discussion and Conclusion

We have identified the contrast effect as a plausible driver of skewness preferences. Accordingly, when comparing a risky and a safe option, a risky outcome attracts the more attention the more it differs from the safe option's payoff. Thereby, the contrast effect induces a focus on the large, but unlikely upside of right-skewed risks, and a focus on the large potential loss in the case of left-skewed risks. As a consequence, salience theory (BGS) and related approaches to local thinking that incorporate contrast effects, such as a model of focusing (KS), predict a preference for positive skewness. In contrast, a model of relative thinking (Bushong *et al.*, 2016), that assumes reverse contrast effects (i.e., the weight assigned to a risky outcome decreases in its contrast to the safe option's payoff), cannot account for skewness preferences.

Models of local thinking offer an explanation for skewness preferences that fundamentally differs from approaches previously proposed in the literature. According to cumulative prospect theory, for instance, an agent exhibits a preference for right-skewed lotteries because she overweights small probabilities *per se*. In contrast, local thinkers overweight a small probability only if the corresponding payoff stands out. This approach to model non-linear probability weighting is not only psychologically sound, but also allows for more realistic predictions. In case of a concave value function, larger payoffs are less attractive and less attention-grabbing so that also the corresponding probabilities should be less distorted, and salience-effects should have less impact on choices. Hence, local thinkers with a value function that is close-to-linear or even convex exhibit strong skewness preferences, and therefore gamble excessively, while local thinkers with a sufficiently concave value function will not do so. As a consequence, local thinking in combination with a population of heterogeneous agents who differ with respect to their intrinsic risk attitudes allows for more plausible predictions on the magnitude of skewness preferences than cumulative prospect theory (Rieger and Wang, 2006; Azevedo and Gottlieb, 2012; Ebert and Strack, 2015, 2016).

Our approach has also advantages compared to other behavioral explanations for skewness preferences such as the model on optimal expectations proposed by Brunnermeier and Parker (2005). Here, an agent receives utility not only from her actions, but also from her beliefs over the likelihood of favorable future outcomes. Therefore, an agent inflates the "perceived likelihood" of upside events in order to enhance the pleasure from expecting these events. As a consequence, the demand for right-skewed lotteries is excessive. This model, however, yields weaker predictions on skewness preferences than our approach (see Proposition 2 in Brunnermeier and Parker, 2005). First, Brunnermeier and Parker explain an affection toward sufficiently right-skewed risks, but they do not obtain precise predictions on the demand for less skewed or left-skewed assets. Second, the puzzle investigated by Ebert and Strack (2015) cannot be resolved in their framework as long as the value function is unbounded. Finally, utility from pleasant expectations can be obtained only *before* an event is realized. Thus, it is plausible that optimal expectations matter only when there is a considerable amount of time between an investment decision and the event realization. Models of local thinking instead can explain skewness preferences irrespective of whether the realization of outcomes is delayed or not.

While a substantial body of evidence suggests that people like right-skewed, but avoid left-skewed risks, we derive a novel set of predictions on skewness-dependent risk attitudes that can be tested in the lab. First, models of local thinking predict that a risky option is preferred over its expected value if and only if its skewness exceeds a certain threshold. Second, the salience model predicts that a risky option with positive payoffs becomes the more attractive the larger its expected value is. While the former prediction has not been tested yet, the latter one has been confirmed in a recent experimental study by Frydman and Mormann (2016).

Beside skewness preferences, models of local thinking can account for a wide range of decision anomalies. In particular the salience model rationalizes empirical observations such as the Allais paradox (Bordalo *et al.*, 2012), decoy effects (Bordalo *et al.*, 2013b), or the newsvendor problem (Dertwinkel-Kalt and Köster, 2017) in one coherent framework, thereby challenging cumulative prospect theory as the major behavioral model of individual decision-making. Its assumptions have been supported both by empirical (Hastings and Shapiro, 2013) and experimental (Dertwinkel-Kalt *et al.*, 2017b) work. Consequently, models of local thinking improve our understanding of when agents seek and when they shy away from risk.

Appendix A: Proofs

A.1: Auxiliary Result

We characterize the ordering property via the partial derivatives of the salience function.

Lemma 2. *Without loss of generality assume $x \geq y$. Consider a continuously differentiable function $\sigma : \mathbb{R}^2 \to \mathbb{R}_+$. Then, the following two statements are equivalent:*

i) σ satisfies ordering.

ii) $\frac{\partial}{\partial x}\sigma(x,y) \geq 0 \geq \frac{\partial}{\partial y}\sigma(x,y)$ for all $(x,y) \in \mathbb{R}^2$ and $\frac{\partial}{\partial x}\sigma(x,y) > 0 > \frac{\partial}{\partial y}\sigma(x,y)$ on any dense subset of $\mathbb{R}$.

Proof. i) $\Rightarrow$ ii): For any $\epsilon > 0$ the ordering property implies $\sigma(x+\epsilon, y) - \sigma(x,y) > 0$. Thus, $\frac{\partial}{\partial x}\sigma(x,y) \geq 0$ follows immediately from the definition of the partial derivative. By the mean value theorem, there exists some $\xi \in [x, x+\epsilon]$ such that

$$\frac{\partial}{\partial x}\sigma(x,y)\bigg|_{x=\xi} = \frac{\sigma(x+\epsilon,y) - \sigma(x,y)}{\epsilon} > 0.$$

Hence, $\frac{\partial}{\partial x}\sigma(x,y) > 0$ on any dense subset of $\mathbb{R}$. The argument for $\frac{\partial}{\partial y}\sigma(x,y)$ is analogous.
ii) $\Rightarrow$ i): For any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$ it holds that

$$\sigma(x+\epsilon, y-\epsilon') - \sigma(x,y) = \int_x^{x+\epsilon} \frac{\partial}{\partial z}\sigma(z, y-\epsilon')dz - \int_{y-\epsilon'}^{y} \frac{\partial}{\partial z}\sigma(x,z)dz > 0$$

since either $[x, x+\epsilon]$ or $[y-\epsilon', y]$ or both are non-empty and dense subsets of $\mathbb{R}$. Hence, σ satisfies ordering. $\square$

A.2: Main Results

Now we prove the results stated in the main text.

Proof of Proposition 1. Consider some discrete lottery $L := (x_1, p_1; \ldots; x_n, p_n)$ with $n \geq 2$. Denote $\Delta_i := \Delta^{-\sigma(u(x_i),u(c))}$ and $\sigma^i := \sigma(u(x_i), u(c))$ as well as $\sigma_x^i := \frac{\partial \sigma^i}{\partial u(x_i)}$ and $\sigma_y^i := \frac{\partial \sigma^i}{\partial u(c)}$.

First we verify that the certainty equivalent is unique. For that, it is sufficient to show

$$\frac{\partial U^s(L)}{\partial u(c)} = -\ln(\Delta)\left(\frac{\left(\sum_{k=1}^n p_k\, u(x_k)\Delta_k\sigma_y^k\right)\left(\sum_{k=1}^n p_k\,\Delta_k\right) - \left(\sum_{k=1}^n p_k\Delta_k\sigma_y^k\right)\left(\sum_{k=1}^n p_k\, u(x_k)\Delta_k\right)}{\left(\sum_{k=1}^n p_k\,\Delta_k\right)^2}\right) \leq 0.$$

Now it is straightforward to see that $\frac{\partial U^s(L)}{\partial u(c)} \leq 0$ holds if and only if

$$\underbrace{\frac{\sum_{k=1}^n p_k\, u(x_k)\Delta_k}{\sum_{k=1}^n p_k\,\Delta_k}}_{=u(c)}\left(\sum_{k=1}^n p_k\Delta_k\sigma_y^k\right) \geq \sum_{k=1}^n p_k\, u(x_k)\Delta_k\sigma_y^k.$$

Denote $\underline{X} := \{k \in \{1,\ldots,n\} | u(x_k) \le u(c)\}$ and $\overline{X} := \{k \in \{1,\ldots,n\} | u(x_k) > u(c)\}$. Then, we can rewrite the above inequality as

$$\sum_{k\in\underline{X}} p_k\, \Delta_k\, \underbrace{\sigma_y^k}_{\ge 0} \underbrace{(u(c)-u(x_k))}_{\ge 0} + \sum_{k\in\overline{X}} p_k\, \Delta_k\, \underbrace{\sigma_y^k}_{\le 0} \underbrace{(u(c)-u(x_k))}_{<0} \ge 0.$$

Hence, $\frac{\partial U^s(L)}{\partial u(c)} \le 0$ always holds and the certainty equivalent is unique.

Second, we verify that the certainty equivalent is monotonic in outcomes. Denote

$$H(\mathbf{x},\mathbf{p},c) := u^{-1}\left(\frac{\sum_{i=1}^n p_i\, u(x_i)\Delta^{-\sigma(u(x_i),u(c))}}{\sum_{i=1}^n p_i\, \Delta^{-\sigma(u(x_i),u(c))}}\right) - c,$$

where $\mathbf{x} := (x_1,\ldots,x_n)$, $\mathbf{p} := (p_1,\ldots,p_n)$. Then, we observe that

$$\frac{\partial}{\partial c} H(\mathbf{x},\mathbf{p},c) = \underbrace{(u^{-1})'\,(U^s(L))}_{>0}\, \underbrace{u'(c)}_{>0}\, \underbrace{\frac{\partial U^s(L)}{\partial u(c)}}_{\le 0} -1 < 0 \tag{C.2}$$

and

$$\frac{\partial}{\partial x_k} H(\mathbf{x},\mathbf{p},c) = \underbrace{(u^{-1})'\,(U^s(L))}_{>0}\, \underbrace{u'(x_k)}_{>0}\, \frac{\partial U^s(L)}{\partial u(x_k)}$$

where

$$\frac{\partial U^s(L)}{\partial u(x_k)} = \frac{[p_k\Delta_k - p_k\Delta_k \ln(\Delta)\sigma_x^k u(x_k)]\left(\sum_{i=1}^n p_i\,\Delta_i\right) - [p_k\Delta_k(-\ln(\Delta))\sigma_x^k]\left(\sum_{i=1}^n p_i\, u(x_i)\Delta_i\right)}{\left(\sum_{i=1}^n p_i\,\Delta_i\right)^2}.$$

Thus, we have $\frac{\partial U^s(L)}{\partial u(x_k)} > 0$ if and only if

$$p_k\Delta_k\left[1 - \ln(\Delta)\sigma_x^k u(x_k)\right] > p_k\Delta_k(-\ln(\Delta))\sigma_x^k \underbrace{\left(\frac{\sum_{i=1}^n p_i\, u(x_i)\Delta_i}{\sum_{i=1}^n p_i\,\Delta_i}\right)}_{=u(c)}$$

or, equivalently,

$$1 + \underbrace{\ln(\Delta)}_{<0}\ \underbrace{\sigma_x^k\left(u(c) - u(x_k)\right)}_{\le 0} > 0.$$

This inequality is always fulfilled as $\sigma_x^k \ge 0$ holds if and only if $u(c) \le u(x_k)$. Hence, we have $\partial H(\mathbf{x},\mathbf{p},c)/\partial x_k > 0$ and the implicit function theorem yields monotonicity in outcomes, that is,

$$\frac{\partial c}{\partial x_k} = -\frac{\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial x_k}}{\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial c}} > 0.$$

Third, we assess whether the certainty equivalent is also monotonic in probabilities. Suppose that probability mass is c.p. shifted from outcome x_l to outcome x_i for some $i, l \in \{1,\ldots,n\}$, $i \ne l$. By definition, a salient thinker's certainty equivalent is monotonic in probabilities if and only if

$$\frac{\partial c}{\partial p_i} > 0 \Leftrightarrow x_i > x_l.$$

Denote $p_l = 1 - \sum_{j \neq l} p_j$ so that an increase in p_i induces a corresponding decrease in p_l. The implicit function theorem yields

$$\frac{\partial c}{\partial p_i} = -\frac{\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial p_i}}{\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial c}}.$$

Using ineq. (C.2) the certainty equivalent is monotonic in probabilities if and only if

$$\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial p_i} > 0 \Leftrightarrow x_i > x_l.$$

Suppose $x_i > x_l$. Then we observe that

$$\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial p_i} = \underbrace{(u^{-1})'\,(U^s(L))}_{>0} \left(\frac{[u(x_i)\ \Delta_i - u(x_l)\ \Delta_l] \sum_{k=1}^{n}(p_k\ \Delta_k) - [\Delta_i - \Delta_l] \sum_{k=1}^{n}(p_k\ u(x_k)\Delta_k)}{\left(\sum_{k=1}^{n} p_k\ \Delta_k\right)^2} \right) > 0,$$

which holds if and only if

$$\big(u(x_i) - u(c)\big)\Delta_i > \big(u(x_l) - u(c)\big)\Delta_l. \tag{C.3}$$

We distinguish the following three cases:

(1) $x_i > x_l > c$: In this case $u(x_i) - u(c) > u(x_l) - u(c) > 0$ and $\Delta_i > \Delta_l$ due to ordering. Thus, (C.3) is satisfied.

(2) $x_i > c > x_l$: The left-hand side of (C.3) is positive, while its right-hand side is negative, so that inequality (C.3) holds.

(3) $c > x_i > x_l$: Here, $0 > u(x_i) - u(c) > u(x_l) - u(c)$ and $\Delta_i < \Delta_l$ due to ordering which gives $(u(x_i) - u(c))\Delta_i > (u(x_i) - u(c))\Delta_l > (u(x_l) - u(c))\Delta_l$.

The case $x_i < x_l$ is analogous. Altogether, we conclude

$$\frac{\partial H(\mathbf{x},\mathbf{p},c)}{\partial p_i} > 0 \quad \text{if and only if} \quad x_i > x_l.$$

This completes the proof. □

Proof of Proposition 2. Let $L := (x_1, p; x_2, 1-p)$ with $x_2 > x_1$. Ordering implies

$$\lim_{p \to 0} \sigma(x_1, \mathbb{E}[L]) = \sigma(x_1, x_2) > \sigma(x_2, x_2) = \lim_{p \to 0} \sigma(x_2, \mathbb{E}[L]).$$

Since the salience function is continuous, there exists some $\hat{p} = \hat{p}(x_1, x_2) \in (0,1)$ such that the lottery's downside is weakly more salient than its upside for any $p \leq \hat{p}$. Then, the statement immediately follows from the fact that—due to ordering—the salience of the lottery's downside payoff x_1 monotonically decreases in the probability p, while the salience of its upside payoff monotonically increases in p. $\square$

Proof of Proposition 3. Consider a binary lottery L with expected value E and variance V. For a given skewness S, its parameters x_1, x_2 and p are uniquely defined as delineated in Lemma 1. Now suppose the lottery's skewness increases. Then, we observe that the lottery's downside payoff becomes more likely. Formally, we have

$$\frac{\partial p}{\partial S} = 2 \cdot (S^2 + 4)^{-3/2} > 0.$$

Using (2), this implies that both the downside payoff x_1 and the upside payoff x_2 increase in the skewness S. Therefore, the difference between the downside (upside) payoff and the expected value decreases (increases) in the lottery's skewness S. Formally, we have

$$\frac{\partial(E - x_1)}{\partial S} < 0 \quad \text{and} \quad \frac{\partial(x_2 - E)}{\partial S} > 0.$$

Since the expected value E is fixed, an increase in contrast is equivalent to an increase in salience due to ordering. Hence, the downside payoff's salience decreases in S, while the upside payoff's salience increases in S.

Since $\lim_{S\to\infty} x_2 = \infty > E$, we obtain

$$\lim_{S\to\infty} \sigma(x_2, E) > \sigma(E, E) = \lim_{S\to\infty} \sigma(x_1, E)$$

by the ordering property. Now by continuity of the salience function we can conclude that there exists some $\hat{S} < \infty$ such that for any $S > \hat{S}$ the lottery's upside is salient. As we have seen that the salience of both outcomes is monotonic in the lottery's skewness S, we conclude that the salient thinker chooses the risky option if and only if $S > \hat{S}$. Finally, $\lim_{S\to-\infty} \sigma(x_1, E) > \sigma(E, E) = \lim_{S\to-\infty} \sigma(x_2, E)$ and monotonicity ensure that there exists a unique skewness value $\hat{S} \in \mathbb{R}$ such that $r(E, V, \hat{S}) = 0$. $\square$

Proof of Proposition 4. Consider two salience functions σ and $\hat{\sigma}$. Suppose that the contrast effect is stronger for salience function σ than for salience function $\hat{\sigma}$. For the binary lottery L with expected value E, variance V, and skewness S, denote $r(E, V, S)$ the risk premium if the salience of outcomes is assessed via σ and $\hat{r}(E, V, S)$ the risk premium if the salience of outcomes is assessed via $\hat{\sigma}$. Then, it holds $r(E, V, S) > \hat{r}(E, V, S)$ if and only if

$$\frac{\sqrt{Vp(1-p)}(\Delta_1 - \Delta_2)}{p\Delta_1 + (1-p)\Delta_2} > \frac{\sqrt{Vp(1-p)}(\hat{\Delta}_1 - \hat{\Delta}_2)}{p\hat{\Delta}_1 + (1-p)\hat{\Delta}_2} \tag{C.5}$$

where $\Delta_k := \Delta^{-\sigma(x_k, E)}$ and $\hat{\Delta}_k := \Delta^{-\hat{\sigma}(x_k, E)}$ for $k \in \{1, 2\}$. Rewriting (C.5) gives

$$\frac{\Delta_1/\Delta_2 - 1}{p\Delta_1/\Delta_2 + (1-p)} > \frac{\hat{\Delta}_1/\hat{\Delta}_2 - 1}{p\hat{\Delta}_1/\hat{\Delta}_2 + (1-p)}$$

or, equivalently,

$$\frac{\Delta_1}{\Delta_2} > \frac{\hat{\Delta}_2}{\hat{\Delta}_2}.$$

Then, applying the definition of salience weights yields

$$\Delta^{-\sigma(x_1,E)+\sigma(x_2,E)} > \Delta^{-\hat{\sigma}(x_1,E)+\hat{\sigma}(x_2,E)},$$

which holds if and only if

$$\sigma(x_2, E) - \sigma(x_1, E) < \hat{\sigma}(x_2, E) - \hat{\sigma}(x_1, E).$$

Rearranging this inequality gives

$$\sigma(x_2, E) - \hat{\sigma}(x_2, E) < \sigma(x_1, E) - \hat{\sigma}(x_1, E).$$

This holds if and only if

$$\sqrt{\frac{Vp}{1-p}} = x_2 - E < E - x_1 = \sqrt{\frac{V(1-p)}{p}} \tag{C.6}$$

since the contrast effect is stronger for σ than for $\hat{\sigma}$. Finally, we conclude that (C.6) holds if and only if $p < 1/2$ or, equivalently, $S < 0$. By Definition 3, this is the case if and only if the lottery $L(E, V, S)$ is left-skewed. □

Proof of Proposition 5. For $x_2 > x_1 \geq 0$, let lottery $L := (x_1, p; x_2, 1-p)$ satisfy Condition (C.1) given salience function $\hat{\sigma}$. Then, it is immediate that the upside of lottery L is salient under salience function $\hat{\sigma}$. As a consequence, it has to hold that

$$u(x_2) - u(\mathbb{E}[L]) > u(\mathbb{E}[L]) - u(x_1). \tag{C.7}$$

To see this, assume the opposite. Then, since $u(x_2) > u(x_1) \geq 0$, we have

$$\begin{aligned}\hat{\sigma}(u(x_1), u(\mathbb{E}[L])) &> \hat{\sigma}(u(\mathbb{E}[L]), u(\mathbb{E}[L]) + u(\mathbb{E}[L]) - u(x_1)) \\ &\geq \hat{\sigma}(u(\mathbb{E}[L]), u(\mathbb{E}[L]) + u(x_2) - u(\mathbb{E}[L])) \\ &= \hat{\sigma}(u(x_2), u(\mathbb{E}[L])),\end{aligned}$$

where the first inequality follows from diminishing sensitivity, the second one from ordering, and the final equality from symmetry. This yields a contradiction to the fact that the upside of lottery L is salient.

From Condition (C.7), we conclude

$$\sigma(u(x_2), u(\mathbb{E}[L])) - \hat{\sigma}(u(x_2), u(\mathbb{E}[L])) > \sigma(u(x_1), u(\mathbb{E}[L])) - \hat{\sigma}(u(x_1), u(\mathbb{E}[L]))$$

by Definition 4 as the contrast effect is stronger for salience function σ than for salience function $\hat{\sigma}$. Rearranging the above inequality yields

$$\sigma(u(x_2), u(\mathbb{E}[L])) - \sigma(u(x_1), u(\mathbb{E}[L])) > \hat{\sigma}(u(x_2), u(\mathbb{E}[L])) - \hat{\sigma}(u(x_1), u(\mathbb{E}[L])).$$

As $\Delta < 1$ and $\hat{\sigma}(u(x_2), u(\mathbb{E}[L])) > \hat{\sigma}(u(x_1), u(\mathbb{E}[L]))$ we conclude

$$\Delta^{\sigma(u(x_2),u(\mathbb{E}[L]))-\sigma(u(x_1),u(\mathbb{E}[L]))} < \Delta^{\hat{\sigma}(u(x_2),u(\mathbb{E}[L]))-\hat{\sigma}(u(x_1),u(\mathbb{E}[L]))}.$$

Thus, if lottery L satisfies Condition (C.1) for salience function $\hat{\sigma}$, then lottery L also satisfies Condition (C.1) for salience function σ. This completes the proof. □

Proof of Proposition 6. For a given expected value $E \in \mathbb{R}$, consider a lottery $L \in \mathcal{L}(E)$ which is sold at some price $z \in \mathbb{R}$. Hence the choice set comprises $\{L, z\}$. As u is (weakly) concave there exist some $a, b \geq 0$ such that $u(x) \leq ax + b$. Denote $\Delta_k := \Delta^{-\sigma(u(x_k),u(z))}$ for $k \in \{1, 2\}$. Using $p = (x_2 - E)/(x_2 - x_1)$ we get

$$\begin{aligned}
U^s(L) =& \frac{\Delta_1(x_2 - E)u(x_1) + \Delta_2(E - x_1)u(x_2)}{\Delta_1(x_2 - E) + \Delta_2(E - x_1)} \\
\leq& \frac{\Delta_1(x_2 - E)(ax_1 + b) + \Delta_2(E - x_1)(ax_2 + b)}{\Delta_1(x_2 - E) + \Delta_2(E - x_1)} \\
=& b + a \cdot \frac{\Delta_1(x_2 - E)x_1 + \Delta_2(E - x_1)x_2}{\Delta_1(x_2 - E) + \Delta_2(E - x_1)} \\
\leq& b + a\bar{\Delta} \cdot \frac{(x_2 - E)x_1 + (E - x_1)x_2}{x_2 - x_1} \\
=& b + a\bar{\Delta}E.
\end{aligned}$$

Here, the first inequality follows from the concavity of the value function, while the second inequality follows from using the upper bound of $\bar{\Delta}$ for the salience weights in the numerator and the lower bound of 1 for the salience weights in the denominator. □

A.3: Comparative Statics of the Salience Model

Here, we derive the comparative statics underlying Prediction 2. More specifically, we delineate how the threshold value $\hat{S}$ (defined in Proposition 3) depends on the lottery's expected value E if we assume a linear value function. In order to derive these comparative statics, we have to impose more structure on the salience function.

Definition 6 (Marginal ordering)**.** *A salience function satisfies the* marginal ordering property *if and only if* $\frac{\partial}{\partial x}\sigma(x, y) > 0 > \frac{\partial}{\partial y}\sigma(x, y)$ *for any* $x > y$.

Note that all salience functions that have been proposed in the literature satisfy the marginal ordering property—in particular, our standard salience function $\sigma_{\beta,\theta}$. In this sense, the assumption of marginal ordering is not very restrictive. Using this property, we are able to prove the following lemma.

Lemma 3. *Suppose the salience function satisfies the marginal ordering property. Then, the threshold value $\hat{S}$ defined in Proposition 3 satisfies*

$$\frac{\partial}{\partial E}\hat{S}(E,V) < 0 \text{ if and only if } \frac{\partial}{\partial E}\Big(\sigma(x_2(E,V,S),E) - \sigma(x_1(E,V,S),E)\Big)\Big|_{S=\hat{S}} > 0.$$

Proof. Consider the lottery $L(E,V,S)$, which has outcomes $x_k = x_k(E,V,S)$, $k \in \{1,2\}$, and a downside probability $p = p(S)$ as defined in (2). We divide the proof into two parts. First, we investigate how the salient thinker's risk premium $r(E,V,S)$ depends on the lottery's expected value E. Second, we use this result to prove our lemma.

PART (1). We need to determine the sign of

$$\frac{\partial}{\partial E}r(E,V,S) = \sqrt{Vp(1-p)} \cdot \frac{\partial}{\partial E}\left(\frac{\Delta^{-\sigma(x_1,E)} - \Delta^{-\sigma(x_2,E)}}{p\Delta^{-\sigma(x_1,E)} + (1-p)\Delta^{-\sigma(x_2,E)}}\right), \tag{4}$$

which equals the sign of

$$\frac{\partial}{\partial E}\left(\frac{\Delta_1 - \Delta_2}{p\Delta_1 + (1-p)\Delta_2}\right) = \frac{\left(\frac{\partial\Delta_1}{\partial E} - \frac{\partial\Delta_2}{\partial E}\right)(p\Delta_1 + (1-p)\Delta_2) - (\Delta_1 - \Delta_2)\left(p\frac{\partial\Delta_1}{\partial E} + (1-p)\frac{\partial\Delta_2}{\partial E}\right)}{(p\Delta_1 + (1-p)\Delta_2)^2} \tag{5}$$

with $\Delta_k := \Delta^{-\sigma(x_k,E)}$ for $k \in \{1,2\}$. Plugging

$$\frac{\partial\Delta_k}{\partial E} = -\ln(\Delta)\Delta_k\frac{\partial}{\partial E}\sigma(x_k,E)$$

into (5) yields

$$\frac{\partial}{\partial E}\left(\frac{\Delta_1 - \Delta_2}{p\Delta_1 + (1-p)\Delta_2}\right) = \underbrace{\frac{-\ln(\Delta)\Delta_1\Delta_2}{(p\Delta_1 + (1-p)\Delta_2)^2}}_{>0} \cdot \left(\frac{\partial}{\partial E}\sigma(x_1,E) - \frac{\partial}{\partial E}\sigma(x_2,E)\right).$$

Hence, we conclude

$$\frac{\partial}{\partial E}r(E,V,S) < 0 \quad \text{if and only if} \quad \frac{\partial}{\partial E}\left(\sigma(x_2,E) - \sigma(x_1,E)\right) > 0.$$

PART (2). By definition, the threshold value $\hat{S} = \hat{S}(E,V)$ solves $r = r(E,V,S) = 0$. Applying the implicit function theorem yields

$$\frac{\partial}{\partial E}\hat{S}(E,V) = -\frac{\frac{\partial}{\partial E}r(E,V,S)}{\frac{\partial}{\partial S}r(E,V,S)}\Bigg|_{S=\hat{S}}.$$

Thus, we have to determine the sign of

$$\frac{\partial r}{\partial S}\Big|_{S=\hat{S}} = \sqrt{V}\underbrace{\left(\frac{\hat{\Delta}_1 - \hat{\Delta}_2}{p\hat{\Delta}_1 + (1-p)\hat{\Delta}_2}\right)}_{=0 \text{ by definition of } \hat{S}} \cdot \frac{\partial}{\partial S}\left(\sqrt{p(1-p)}\right)\Big|_{S=\hat{S}} + \sqrt{Vp(1-p)} \cdot \underbrace{\frac{\partial}{\partial S}\left(\frac{\Delta_1 - \Delta_2}{p\Delta_1 + (1-p)\Delta_2}\right)\Big|_{S=\hat{S}}}_{=:\Psi(E,V,\hat{S})}$$

where $\hat{\Delta}_k := \Delta^{-\sigma(x_k(E,V,\hat{S}),E)}$ for $k \in \{1,2\}$. Hence, we only need to derive the sign of

$$\Psi(E,V,\hat{S}) = \frac{\left(p\hat{\Delta}_1 + (1-p)\hat{\Delta}_2\right)\left(\frac{\partial\Delta_1}{\partial S} - \frac{\partial\Delta_1}{\partial S}\right)\Big|_{S=\hat{S}} + \overbrace{\left(\hat{\Delta}_1 - \hat{\Delta}_2\right)}^{=0 \text{ by def. of } \hat{S}}\left(p\frac{\partial\Delta_1}{\partial S} + (1-p)\frac{\partial\Delta_2}{\partial S} + \frac{\partial p}{\partial S}\left(\Delta_1 - \Delta_2\right)\right)\Big|_{S=\hat{S}}}{\left(p\hat{\Delta}_1 + (1-p)\hat{\Delta}_2\right)^2}$$

which is equivalent to the sign of

$$\left(\frac{\partial\Delta_1}{\partial S} - \frac{\partial\Delta_2}{\partial S}\right)\Big|_{S=\hat{S}} = -\ln(\Delta)\hat{\Delta}_1\left(\frac{\partial}{\partial S}\sigma(x_1,E) - \frac{\partial}{\partial S}\sigma(x_2,E)\right)\Big|_{S=\hat{S}}.$$

We conclude

$$\frac{\partial}{\partial S}\sigma(x_1,E)\Big|_{S=\hat{S}} = -\frac{\sqrt{V}}{2\sqrt{\left(\frac{1-p}{p}\right)}}\frac{\partial\left(\frac{1-p}{p}\right)}{\partial S}\frac{\partial}{\partial x_1}\sigma(x_1,E)\Big|_{S=\hat{S}}$$

and

$$\frac{\partial}{\partial S}\sigma(x_2,E)\Big|_{S=\hat{S}} = \frac{\sqrt{V}}{2\sqrt{\left(\frac{p}{1-p}\right)}}\frac{\partial\left(\frac{p}{1-p}\right)}{\partial S}\frac{\partial}{\partial x_2}\sigma(x_2,E)\Big|_{S=\hat{S}}.$$

Altogether, we have

$$\left(\frac{\partial\Delta_1}{\partial S} - \frac{\partial\Delta_2}{\partial S}\right)\Big|_{S=\hat{S}} = -\ln(\Delta)\hat{\Delta}_1\sqrt{V}\frac{\partial\left(\frac{p}{1-p}\right)}{\partial S}\left(\sqrt{\frac{p}{1-p}}\overbrace{\frac{\partial}{\partial x_1}\sigma(x_1,E)}^{<0} - \sqrt{\frac{1-p}{p}}\overbrace{\frac{\partial}{\partial x_2}\sigma(x_2,E)}^{>0}\right)\Big|_{S=\hat{S}} < 0$$

since σ satisfies the marginal ordering property and $\frac{\partial\left(\frac{p}{1-p}\right)}{\partial S} = -\frac{\partial\left(\frac{1-p}{p}\right)}{\partial S} > 0$. Hence we conclude that $\frac{\partial}{\partial S}r(E,V,S)\big|_{S=\hat{S}} < 0$.

Using PART (1), we obtain $\frac{\partial}{\partial E}\hat{S}(E,V) < 0$ if and only if

$$\frac{\partial}{\partial E}\left(\sigma(x_2(E,V,S),E) - \sigma(x_1(E,V,S),E)\right)\Big|_{S=\hat{S}} > 0 \tag{C.8}$$

by the implicit function theorem. □

Proposition 7. *Suppose the salience function satisfies the marginal ordering property and decreasing diminishing sensitivity. Then, the threshold value $\hat{S}$ defined in Propostion 3 satisfies*

$$\frac{\partial}{\partial E}\hat{S}(E,V)\begin{cases} > 0, & \text{if } \hat{x}_1 < 0 < \hat{x}_2, \\ < 0, & \text{otherwise,}\end{cases}$$

where $\hat{x}_k := x_k(E,V,\hat{S})$, $k \in \{1,2\}$, is given in Equation (2).

Proof. First, we consider the case $\hat{x}_1 < 0 < \hat{x}_2$. Then, diminishing sensitivity implies that

$$\frac{\partial}{\partial E}\sigma(x_2(E,V,S),E)\Big|_{S=\hat{S}} < 0 < \frac{\partial}{\partial E}\sigma(x_1(E,V,S),E)\Big|_{S=\hat{S}},$$

and the statement follows immediately from the proof of Lemma 3.

Second, we consider the cases where either $\hat{x}_1 \geq 0$ or $\hat{x}_2 \leq 0$. By definition of $\hat{S}$, we have $\sigma(\hat{x}_1, E) = \sigma(\hat{x}_2, E)$. Thus, decreasing diminishing sensitivity implies

$$\frac{\partial}{\partial E}\left(\sigma(x_2(E,V,S),E) - \sigma(x_1(E,V,S),E)\right)\Big|_{S=\hat{S}} > 0,$$

and the statement follows from Lemma 3. □

We have investigated how the threshold value $\hat{S}$ defined in Proposition 3 depends on the expected value of the risk at hand. An immediate consequence of Proposition 7 is Prediction 2, which is stated in the main text.

Finally, it is straigthforward to see that a wide class of salience functions satisfies decreasing diminishing sensitvity, and therefore yields Proposition 7. For any $n \in \mathbb{N}$, for instance, the salience function $\sigma(x,y) = (\sigma_{\beta,\theta}(x,y))^n$ satisfies decreasing diminishing sensitivity. Fix $x, y, z > 0$ such that $x \geq z$ and $\sigma(x-z,x) \geq \sigma(x+y,x)$. Then, we have

$$\frac{\partial}{\partial x}\left[(\sigma_{\beta,\theta}(x-z,x))^n - (\sigma_{\beta,\theta}(x+y,x))^n\right] = 4n\cdot\left[-\frac{(\sigma_{\beta,\theta}(x-z,x))^n}{2x-z+\theta} + \frac{(\sigma_{\beta,\theta}(x+y,x))^n}{2x+y+\theta}\right] < 0,$$

as well as

$$\begin{aligned}\frac{\partial}{\partial x}&\left[(\sigma_{\beta,\theta}(-x+z,-x))^n - (\sigma_{\beta,\theta}(-x-y,-x))^n\right] \\ &= 4n\cdot\left[-\frac{(\sigma_{\beta,\theta}(-x+z,-x))^n}{2x-z+\theta} + \frac{(\sigma_{\beta,\theta}(-x-y,-x))^n}{2x+y+\theta}\right] < 0.\end{aligned}$$

Note that, by the reflection property, it follows $\sigma(-x+z,-x) \geq \sigma(-x-y,-x)$ from $\sigma(x-z,x) \geq \sigma(x+y,x)$. Thus, the salience function $\sigma(x,y)$ satisfies decreasing diminishing sensitivity. Similiar calculations as above show that also the salience function $\sigma(x,y) = \frac{|x-y|}{|x|+|y|+\theta}$, which has been proposed by Bordalo *et al.* (2012), satisfies decreasing diminishing sensitivity and yields the same prediction. To the best of our knowledge, Proposition 7 applies to any salience function proposed in the literature.

Appendix B: Skewness Preferences According to a Model of Focusing (Kőszegi and Szeidl, 2013)

In this section, we verify that our explanation for skewness preferences does not hinge on the specific assumptions of salience theory of choice under risk, but also holds under a related approach to stimulus driven attention—*a model of focusing* (Kőszegi and Szeidl, 2013). As Kőszegi and Szeidl analyze deterministic choice problems only, we extend their model toward risky choices along the lines of salience theory, that is, the agent evaluates an option according to the underlying state space. As discussed in Section 2, this assumption can be relaxed for both the salience and the focusing model.

Model. Suppose some choice set $C := \{L_x, L_y\}$ where $L_x := (x_1, p_1; \ldots, x_n; p_n)$ and $L_y := (y_1, q_1; \ldots; y_m, q_m)$ are discrete lotteries with $n, m \in \mathbb{N}$ and $\sum_{i=1}^{n} p_i = \sum_{i=1}^{m} q_i = 1$. We impose the same conventions for the lotteries' outcomes as in the main text (i.e., the lotteries' ouctomes are pairwisely distinct and occur with strictly positive probability). The state space S comprises all feasible payoff-combinations of the available lotteries. Thereby, each state of the world $s_{ij} := (x_i, y_j)$ occurs with some objective probability π_{ij}. Again we assume that the decision-maker evaluates monetary outcomes via a strictly increasing value function $u(\cdot)$ with $u(0) = 0$.

According to the focusing model, a decision-maker assigns a weight to each state s_{ij} that depends on the state's objective probability π_{ij} and on the absolute difference in the values of the feasible outcomes in this state, denoted as $d_{ij} := |u(x_i) - u(y_j)|$. The larger the range of values assigned to the outcomes in a state is, the higher the agent's focus on this particular state. Formally, the agent's focus on state $s_{ij} \in S$ is given by $g(d_{ij})$ where the focusing function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is bounded and strictly increasing.[19]

For reasons of comparability, we adopt the smooth salience characterization introduced in Section 2 for the focusing model. That is, each state s_{ij} receives focus weight $\Delta^{-g(d_{ij})}$ for some focusing function $g(\cdot)$ and some constant $\Delta \in (0, 1]$ that captures the agent's susceptibility to focusing. We call an agent with $\Delta < 1$ a *focused thinker*.

Definition 7. *A focused thinker's decision utility $U^f(\cdot)$ for $L_x \in \{L_x, L_y\}$ is given by*

$$U^f(L_x) = \sum_{s_{ij} \in S} \pi_{ij} u(x_i) \cdot \frac{\Delta^{-g(d_{ij})}}{\sum_{s_{ij} \in S} \pi_{ij} \Delta^{-g(d_{ij})}}.$$

The normalization factor in the denominator ensures that the distorted probabilities sum up to one and that the valuation for a safe option $c \in \mathbb{R}$ is undistorted; that is, irrespective of the composition of the choice set we have $U^f(c) = U(c) = u(c)$.

Certainty equivalents and monotonicity. Suppose the agent faces some choice set $\{L, c\}$ where $L := (x_1, p_1; \ldots; x_n, p_n)$ is a lottery with $n \geq 2$ pairwisely distinct ouctomes and c denotes the option that pays an amount of $c \in \mathbb{R}$ with certainty. A focused thinker (weakly) prefers the lottery L over the safe option c if and only if

$$U^f(c) \leq U^f(L) = \frac{\sum_{i=1}^{n} p_i u(x_i) \Delta^{-g(|u(x_i) - u(c)|)}}{\sum_{i=1}^{n} p_i \Delta^{-g(|u(x_i) - u(c)|)}} =: F(c).$$

Without loss of generality we assume $x_1 < \ldots < x_n$. Then a focused thinker's certainty equivalent is implicitly given by $c = u^{-1}(F(c))$. As for the salience model, we conclude that $u^{-1} \circ F : [x_1, x_n] \to [x_1, x_n]$ has at least one fixed point due to Brouwer's fixed-point theorem and we obtain the following proposition.

[19] Relatedly, Bushong *et al.* (2016) propose *a model of relative thinking* that differs from the preceding focusing model only in the assumption on the slope of g: while we have $g'(d_{ij}) > 0$ for the focusing model, we have $g'(d_{ij}) < 0$ for the model of relative thinking. In words, a relative thinker's probability weight on state s_{ij} decreases in the corresponding absolute difference in values d_{ij}.

Proposition 8 (Certainty equivalent to a discrete lottery). *A focused thinker's certainty equivalent to a lottery with $n \geq 2$ outcomes is unique and monotonic in outcomes and probabilities.*

Proof. Note that for any salience function $\sigma(\cdot,\cdot)$ and any focusing function $g(\cdot)$ we have

$$sgn\left(\frac{\partial\sigma(u(x_i),u(c))}{\partial u(x_i)}\right) = sgn\left(\frac{\partial g(|u(x_i)-u(c)|)}{\partial u(x_i)}\right) \quad \text{and} \quad sgn\left(\frac{\partial\sigma(u(x_i),u(c))}{\partial u(c)}\right) = sgn\left(\frac{\partial g(|u(x_i)-u(c)|)}{\partial u(c)}\right).$$

Then, the statement simply follows from replacing the salience function in the proof of Proposition 1 by a focusing function. □

Skewness preferences under a linear value function. To investigate a focused thinker's attitude toward skewness, suppose some choice set $\{L, \mathbb{E}[L]\}$ where $L := (x_1, p; x_2, 1-p)$ is a binary lottery with outcomes $x_2 > x_1$ and the expected value $\mathbb{E}[L] := p \cdot x_1 + (1-p) \cdot x_2$. As in Section 4, we assume a linear value function $u(x) = x$.

Using the characterization of binary risks in Lemma 1, a focused thinker's risk premium for the binary lottery L with expected value E, variance V, and skewness S equals

$$r(E,V,S) = \sqrt{Vp(1-p)} \cdot \left(\frac{\Delta^{-g(E-x_1)} - \Delta^{-g(x_2-E)}}{p\Delta^{-g(E-x_1)} + (1-p)\Delta^{-g(x_2-E)}}\right) \tag{6}$$

where outcomes $x_k = x_k(E,V,S)$, $k \in \{1,2\}$, and probability $p = p(S)$ are defined in (2). A focused thinker strictly prefers the lottery $L(E,V,S)$ over the safe option E if and only if the lottery's risk premium is strictly negative, or, equivalently, the agent's focus lies on the lottery's upside payoff. We conclude:

Proposition 9 (Skewness preferences). *For any given expected value E and variance V, a focused thinker strictly prefers the lottery $L(E,V,S)$ over its expected value E if and only if $S > 0$.*

Proof. It is straightforward to show that a focused thinker's risk premium is strictly negative if and only if $g(x_2 - E) > g(E - x_1)$; that is, her focus lies on the lottery's upside payoff. As g is a strictly increasing function, this is the case if and only if

$$\sqrt{V}\sqrt{\frac{p}{1-p}} = x_2 - E > E - x_1 = \sqrt{V}\sqrt{\frac{1-p}{p}},$$

or equivalently, $p > 1/2$. Then, from equation (1), we conclude that a focused thinker strictly prefers the lottery over its expected value if and only if $S > 0$. □

Hence, a focused thinker seeks right-skewed but avoids left-skewed risks.[20]Similar to the salience model, we observe that a focused thinker's preference for right-skewed and aversion toward left-skewed risks is enhanced if the contrast effect becomes stronger.

[20]Note that, for any expected value E and variance V, a *relative thinker* (Bushong *et al.*, 2016, see also footnote 12) prefers the binary lottery $L(E,V,S)$ over its expected value E if and only if $S < 0$. It is straightforward to show that, for a relative thinker, we have $g(x_2 - E) > g(E - x_1)$ if and only if $p < 1/2$ as g is strictly decreasing by assumption. Hence a relative thinker seeks left-skewed but avoids right-skewed risks.

Definition 8. *We say that the contrast effect is stronger for focusing function g than for focusing function $\hat{g}$ if the difference $g(x) - \hat{g}(x)$ is increasing in $x \in \mathbb{R}_+$.*

Note that the argument of the focusing function represents the difference in values assigned to the outcomes that are feasible in a given state. Thus, the preceding definition of the strength of the contrast effect is analogous to the definition given for the salience model. We conclude:

Proposition 10 (Contrast and skewness preferences)**.** *Let the contrast effect be stronger for focusing function g than for focusing function $\hat{g}$. Then, a focused thinker's risk premium $r(E, V, S)$ is larger for g than for $\hat{g}$ if and only if $S < 0$, that is, the lottery is left-skewed.*

Proof. Analogous to the proof of Proposition 4. □

Puzzles on skewness preferences. Similar to salience theory of choice under risk, the focusing approach yields more reasonable predictions on the magnitude of skewness preferences than cumulative prospect theory. We will show that the puzzles on skewness preferences in the small (Ebert and Strack, 2015, 2016) and in the large (Rieger and Wang, 2006; Azevedo and Gottlieb, 2012) arising for CPT agents can be resolved for focused thinkers.

First, we argue that focusing does not necessarily yield the same unrealistic predictions on skewness preferences in the small as cumulative prospect theory (Ebert and Strack, 2015). Formally, suppose that a focused thinker faces some choice set $\{L, \mathbb{E}[L]\}$. In line with Section 5, let $x_2 > x_1 \geq 0$ and assume that the decision-maker's value from money is strictly increasing and strictly concave, that is, $u'(\cdot) > 0$ and $u''(\cdot) < 0$. As before, we normalize $u(0) = 0$. Then, a focused thinker strictly prefers the risky lottery L over the safe option $\mathbb{E}[L]$ if and only if

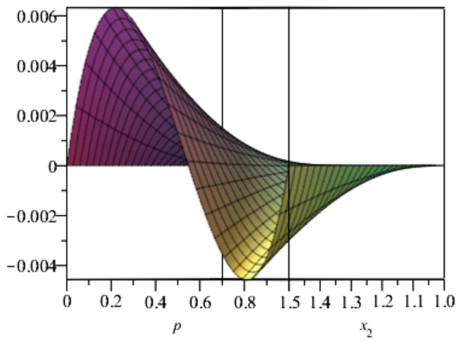
$$\frac{u(x_2) - u(\mathbb{E}[L])}{u(\mathbb{E}[L]) - u(x_1)} \cdot \frac{1-p}{p} > \frac{\Delta_1}{\Delta_2},$$

where $\Delta_k := \Delta^{-g(|u(x_k) - u(\mathbb{E}[L])|)}$, $k \in \{1, 2\}$. For any given expected value $E = \mathbb{E}[L]$, substituting $p = (x_2 - E)/(x_2 - x_1)$ yields

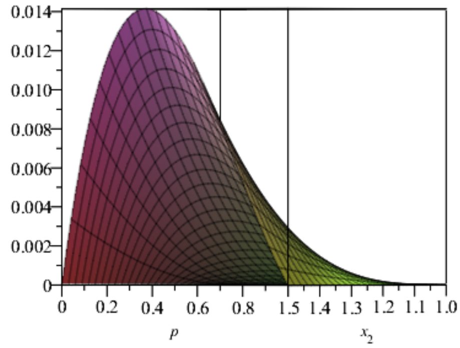
$$\frac{\frac{u(x_2)-u(E)}{x_2-E}}{\frac{u(E)-u(x_1)}{E-x_1}} > \frac{\Delta_1}{\Delta_2}. \tag{C.1–Focus}$$

Using the following two examples, we show that depending on the value function's curvature there might not exist a binary lottery satisfying Condition (C.1–Focus). As in Section 5, we assume power utility $u(x) = x^\alpha$ for some $\alpha \in (0, 1)$. Furthermore, we consider the focusing function $g(x) = 1 - \frac{1}{1+\gamma x}$ for some $\gamma > 0$ and $x \in \mathbb{R}_+$. Let $\Delta = 0.7$.

Example 7 (Value function). Assume $\gamma = 1$ so that the focusing function is given by $g(x) = 1 - \frac{1}{1+x}$. If the value function's curvature increases, that is, the parameter α decreases, we observe that inequality (C.1–Focus) is less likely to hold. More specifically, numerical computations show that there exists some threshold value $\tilde{\alpha} \in (0,1)$ such that for any $\alpha \in (0, \tilde{\alpha})$ no unfair, attractive gamble exists. For $\alpha = 0.95$ and $\alpha = 0.5$, Figure 3 illustrates the risk premium r as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.



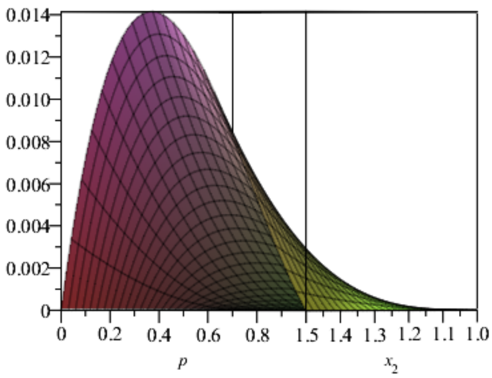
Risk premium for $\alpha = 0.95$.



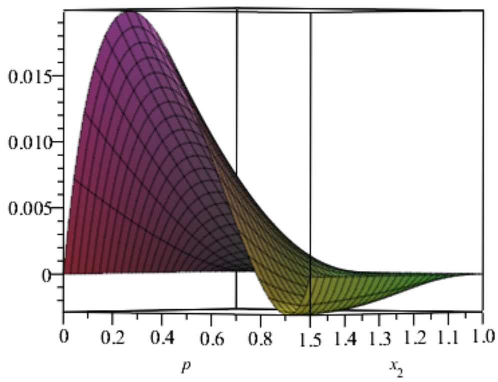
Risk premium for $\alpha = 0.5$.

Figure 3: *The above graphs show the risk premium as a function of the upside payoff x_2 and the probability p that the downside payoff x_1 is realized. For $\alpha = 0.95$, the risk premium becomes negative for highly right-skewed lotteries (i.e., a large probability p on the downside payoff x_1) with a large upside payoff x_2. For $\alpha = 0.5$, the risk premium is non-negative for any feasible lottery.*

Example 8 (Focusing function). Fix $\alpha = 1/2$ so that the value function is $u(x) = \sqrt{x}$. We observe that inequality (C.1–Focus) is more likely to hold for at least some binary lottery L if parameter γ increases, that is, the contrast effect becomes stronger. In fact, numerical computations show that there exists some $\hat{\gamma} > 1$ such that for any $\gamma > \hat{\gamma}$ at least one unfair, attractive gamble exists. For $\gamma = 1$ and $\gamma = 10$, Figure 4 illustrates the risk premium r as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.



Risk premium for $\gamma = 1$.



Risk premium for $\gamma = 10$.

Figure 4: *The above graphs show the risk premium as a function of the upside payoff x_2 and the probability p that the downside payoff x_1 is realized. For $\gamma = 10$, the risk premium becomes negative for highly right-skewed lotteries (i.e., a large probability p on the downside payoff x_1) with a large upside payoff x_2. For $\gamma = 1$, the risk premium is non-negative for any feasible lottery.*

Second, we show that a focused thinker's valuation for binary lotteries with a given expected value $E < \infty$ is bounded. Hence, cumulative prospect theory's predictions on skewness preferences in the large—as delineated by Rieger and Wang (2006) and Azevedo and Gottlieb (2012)—can be resolved in the focusing model.

Proposition 11. *Let $\mathcal{L}(E)$ denote the set of binary lotteries L with finite expected value $E \in \mathbb{R}$. A focused thinker's valuation for some $L \in \mathcal{L}(E)$ is bounded by a function which is affine in E.*

Proof. Since the focusing function is bounded there exists some threshold value $\tilde{\Delta} < \infty$ such that $\Delta^{-g(x)} < \tilde{\Delta}$ for any $x \in \mathbb{R}_+$. The remainder of the proof is analogous to the proof of Proposition 6. □

Appendix C: Mao's Lotteries and Skewness Preferences

Suppose choice set $C := \{L_x, L_y\}$ where $L_x := (x_1, p; x_2, 1-p)$ and $L_y := (y_1, q; y_2, 1-q)$ with outcomes $x_2 > x_1$ and $y_2 > y_1$ and probabilities $p, q \in (0,1)$. As in Section 4, we assume a linear value function $u(x) = x$. Mao (1970) introduced the following class of binary lotteries that allow us to identify skewness preferences.

Definition 9. *Let $p \in \left(0, \frac{1}{2}\right)$. Two perfectly correlated, binary lotteries $L_x := (x_1, p; x_2, 1-p)$ and $L_y := (y_1, 1-p; y_2, p)$ denote a* Mao pair *if both have the same expected value and variance.*

Mao lotteries differ only in their skewness: L_x is left-skewed as its high payoff x_2 occurs with a high probability while lottery L_y is right-skewed as its high payoff y_2 occurs with a small probability (for a formal proof see Ebert and Wiesen, 2011). In line with Definition 3, Ebert and Wiesen (2011) state that "an individual is said to be *skewness seeking* if, for any given Mao pair, she prefers L_y over L_x."

Proposition 12. *For any given Mao pair, a salient thinker prefers L_y over L_x.*

Proof. Due to the perfect correlation of the lotteries, the state space S comprises only two states, that is, $S = \{(x_1, y_2), (x_2, y_1)\}$. Hence a salient thinker prefers the right-skewed lottery L_y over the left-skewed lottery L_x if and only if

$$U^s(L_y) - U^s(L_x) = p(y_2 - x_1)\Delta^{-\sigma(x_1, y_2)} + (1-p)(y_1 - x_2)\Delta^{-\sigma(x_2, y_1)} > 0.$$

Since $p(y_2 - x_1) = -(1-p)(y_1 - x_2) > 0$ by definition—remember that both lotteries have the same expected values—the above inequality simplifies to $\sigma(x_1, y_2) > \sigma(x_2, y_1)$. As $p < 1/2$ by Definition 9, Lemma 1 yields

$$x_1 < y_1 < x_2 < y_2.$$

Thus, ordering implies $\sigma(x_1, y_2) > \sigma(x_2, y_1)$, which was to be proven. □

Finally, it is straightforward to see from the proof above that also a focused thinker prefers L_y over L_x for any given Mao pair. In contrast, a relative thinker in the spirit of Bushong *et al.* (2016) would choose L_x over L_y for any given Mao pair.

References

Åstebro, T., Mata, J. and Santos-Pinto, L. (2015). Skewness seeking: Risk loving, optimism or overweighting of small probabilities? *Theory and Decision*, **78** (2), 189–208.

Azevedo, E. M. and Gottlieb, D. (2012). Risk-neutral firms can extract unbounded profits from consumers with prospect theory preferences. *Journal of Economic Theory*, **147** (3), 1291–1299.

Bali, T. G., Cakici, N. and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, **99** (2), 427–446.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, **27** (1), 173–195.

Barseghyan, L., Molinari, F., O'Donoghue, T. and Teitelbaum, J. C. (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, **103** (6), 2499–2529.

Berkhout, P., Hartog, J. and Webbink, D. (2010). Compensation for earnings risk under worker heterogeneity. *Southern Economic Journal*, **76** (3), 762–790.

Bordalo, P., Nicola Gennaioli and Shleifer, A. (2012). Salience theory of choice under risk. *Quarterly Journal of Economics*, **127** (3), 1243–1285.

—, — and Shleifer, A. (2013a). Salience and asset prices. *American Economic Review, Papers & Proceedings*, **103** (3), 623–628.

—, — and Shleifer, A. (2013b). Salience and consumer choice. *Journal of Political Economy*, **121** (5), 803–843.

Boyer, B., Mitton, T. and Vorkink, K. (2010). Expected idiosyncratic skewness. *Review of Financial Studies*, **23** (1), 169–202.

Brünner, T., Levínskỳ, R. and Qiu, J. (2011). Preferences for skewness: evidence from a binary choice experiment. *European Journal of Finance*, **17** (7), 525–538.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, **95** (4), 1092–1118.

Bushong, B., Rabin, M. and Schwartzstein, J. (2016). A model of relative thinking. `http://people.hbs.edu/jschwartzstein/RelativeThinking.pdf`, unpublished.

Chen, J., Hong, H. and Stein, J. C. (2001). Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics*, **61** (3), 345–381.

CHOI, D., LOU, D. and MUKHERJEE, A. (2016). The effect of superstar firms on college major choice. http://www.pbcsf.tsinghua.edu.cn/research/loudong/paper/W%20The%20Effect%20of%20Superstar%20Firms%20on.pdf, unpublished.

CHUNHACHINDA, P., DANDAPANI, K., HAMID, S. and PRAKASH, A. J. (1997). Portfolio selection and skewness: Evidence from international stock markets. *Journal of Banking & Finance*, **21** (2), 143–167.

CICCHETTI, C. J. and DUBIN, J. A. (1994). A microeconometric analysis of risk aversion and the decision to self-insure. *Journal of Political Economy*, **102** (1), 169–186.

CONRAD, J., DITTMAR, R. F. and GHYSELS, E. (2013). Ex ante skewness and expected stock returns. *Journal of Finance*, **68** (1), 85–124.

DERTWINKEL-KALT, M., GERHARDT, H., RIENER, G., SCHWERTER, F. and STRANG, L. (2017a). Concentration bias in intertemporal choice. https://www.dropbox.com/s/dv20mcu0qkygmjz/Concentration_Bias_in_Intertemporal_Choice.pdf?dl=0, unpublished.

— and KÖSTER, M. (2015). Violations of first-order stochastic dominance as salience effects. *Journal of Behavioral and Experimental Economics*, **59**, 42–46.

— and KÖSTER, M. (2017). Salient compromises in the newsvendor game. *Journal of Economic Behavior & Organization, forthcoming*.

—, LANGE, M., KÖHLER, K. and WENZEL, T. (2017b). Demand shifts due to salience effects: Experimental evidence. *Journal of the European Economic Association*, **15** (3), 626–653.

EBERT, S. (2015). On skewed risks in economic models and experiments. *Journal of Economic Behavior & Organization*, **112**, 85–97.

— and HILPERT, C. (2016). Skewness preference and the popularity of technical analysis. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2354962, unpublished.

— and STRACK, P. (2015). Until the bitter end: On prospect theory in a dynamic context. *American Economic Review*, **105** (4), 1618–1633.

— and — (2016). Never, ever getting started: On prospect theory without commitment. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765550, unpublished.

— and WIESEN, D. (2011). Testing for prudence and skewness seeking. *Management Science*, **57** (7), 1334–1349.

FAMA, E. F. and FRENCH, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, **47** (2), 427–465.

FORREST, D., SIMMONS, R. and CHESTERS, N. (2002). Buying a dream: Alternative models of demand for lotto. *Economic Inquiry*, **40** (3), 485–496.

Frydman, C. and Mormann, M. M. (2016). The role of salience and attention in choice under risk: An experimental investigation. https://finance.uni-mannheim.de/fileadmin/files/areafinance/files/FSS_2017/Frydman_2016_RoleOfSalienceAndAttention.pdf, unpublished.

— and — (2017). The role of salience and attention in choice under risk: An experimental investigation. https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2778822, unpublished.

Garrett, T. A. and Sobel, R. S. (1999). Gamblers favor skewness, not risk: Further evidence from united states' lottery games. *Economics Letters*, **63** (1), 85–90.

Gennaioli, N. and Shleifer, A. (2010). What comes to mind. *Quarterly Journal of Economics*, **125** (4), 1399–1433.

Golec, J. and Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, **106** (1), 205–225.

Green, T. C. and Hwang, B.-H. (2012). Initial public offerings as lotteries: Skewness preference and first-day returns. *Management Science*, **58** (2), 432–444.

Grossman, P. J. and Eckel, C. C. (2015). Loving the long shot: Risk taking with skewed lotteries. *Journal of Risk and Uncertainty*, **51** (3), 195–217.

Hartog, J. and Vijverberg, W. P. (2007). On compensation for risk aversion and skewness affection in wages. *Labour Economics*, **14** (6), 938–956.

Hastings, J. S. and Shapiro, J. M. (2013). Fungibility and consumer choice: Evidence from commodity price shocks. *Quarterly Journal of Economics*, **128** (4), 1449.

Kőszegi, B. and Szeidl, A. (2013). A model of focusing in economic choice. *Quarterly Journal of Economics*, **128** (1), 53–104.

Kontek, K. (2016). A critical note on salience theory of choice under risk. *Economics Letters*, **149**, 168–171.

Mao, J. C. (1970). Survey of capital budgeting: Theory and practice. *Journal of Finance*, **25** (2), 349–360.

Mitton, T. and Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies*, **20** (4), 1255–1288.

Prakash, A. J., Chang, C.-H. and Pactwa, T. E. (2003). Selecting a portfolio with skewness: Recent evidence from us, european, and latin american equity markets. *Journal of Banking & Finance*, **27** (7), 1375–1390.

Rieger, M. O. and Wang, M. (2006). Cumulative prospect theory and the St. petersburg paradox. *Economic Theory*, **28** (3), 665–679.

Sydnor, J. (2010). (Over)insuring modest risks. *American Economic Journal: Applied Economics*, **2** (4), 177–199.

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