

Dertwinkel-Kalt, Markus; Köster, Mats

Conference Paper

Local Thinking and Skewness Preferences

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Decision Theory, No. F09-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Dertwinkel-Kalt, Markus; Köster, Mats (2017) : Local Thinking and Skewness Preferences, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Decision Theory, No. F09-V3, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/168303>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Local Thinking and Skewness Preferences*

Markus Dertwinkel-Kalt^{1,†} Mats Köster^{2,‡}

¹University of Cologne

²Düsseldorf Institute for Competition Economics (DICE)

February 28, 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. Therefore, we 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 it depends on the skewness of the underlying probability distribution whether an agent seeks or avoids a specific risk. As 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; Skewness Preferences.

*We are grateful to Paul Heidhues and Daniel Wiesen for their helpful suggestions.

[†]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 risk averse) and gamble in casinos (i.e., behave 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 occur rarely (e.g., Sydnor, 2010; Barseghyan *et al.*, 2013). For example natural disasters belong to this group of left-skewed risks. At the same time, individuals seek right-skewed risks such as casino gambles 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 (e.g., Ebert and Wiesen, 2011; Ebert, 2015) is denoted as *skewness preferences*.

A compelling explanation for skewness preferences is still missing. Expected utility theory (EUT) predicts either risk averse or risk seeking behavior as it implies a valuation for risky options that is linear in probabilities. Thus, it cannot account for risk attitudes that depend on the skewness of a given probability distribution. In order to match experimental and empirical evidence, cumulative prospect theory (CPT; Tversky and Kahneman, 1992) has proposed non-linear probability weighting. As a CPT agent overweights small probabilities, she exhibits a preference for right-skewed and an aversion toward left-skewed risks. This mechanism, however, does not offer any psychologically sound explanation 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 conclusively explain 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 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 much attention. Overweighting the probability of winning the (salient) jackpot, a local thinker behaves 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 risk averse. Importantly, this line of argumentation holds for different models of stimulus-driven attention, that are, *salience theory of choice under risk* (Bordalo *et al.*, 2012, henceforth: BGS) and a *model*

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

of focusing (Kőszegi and Szeidl, 2013, henceforth: KS). Accordingly, 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 (i) certainty equivalents may not exist and (ii) 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 parameter $\delta < 1$. Then, monotonicity in probabilities may be violated if probability mass is shifted from a low, salient outcome to a larger, but less salient outcome which is strongly discounted. BGS use the simplified, discrete version of their model for analytical ease as long as the use of the underlying continuous model is not necessary. We show that all problems raised by Kontek (2016) are resolved in the continuous salience and focusing models.

Second, we show that models of stimulus-driven attention predict skewness preferences. For the discrete salience model, Bordalo *et al.* (2013a) have argued that salient thinkers like right-skewed and dislike left-skewed assets, but they have not precisely disentangled a salient thinker's preferences for risk and skewness. In contrast, we derive skewness preferences formally that is, we show that a salient thinker is more likely to choose a binary risk if it is *ceteris paribus* (i.e., for given expected value and variance) skewed further to the right. In addition, we single out the channel (*contrast effects*) through which the salience model predicts 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 behavior, never starts to gamble (Ebert and Strack, 2015, 2016). In addition, gambles with an arbitrarily large expected loss may attract CPT agents (Rieger and Wang, 2006; Azevedo and Gottlieb, 2012). Models of stimulus-driven attention do not necessarily share these extreme predictions.

Skewness preferences are relevant in several important economic, especially 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 is below that of comparable stocks. This can be explained by the fact that stocks that conduct an initial public offering 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 is the long-term average return of an initial public offering stock. 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). They further allow 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 labour economics as workers reveal skewness affection in wages (Hartog and Vijverberg, 2007; Berkhout *et al.*, 2010). Altogether, skewness preferences help 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 provide 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 a 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 . Each state of the world corresponds to a payoff-combination of the available lotteries. Suppose an agent chooses a lottery from the set $C := \{L_x, L_y\}$ where $L_x := (x_1, p_1; \dots; x_n, p_n)$ and $L_y := (y_1, q_1; \dots; y_m, q_m)$ with $n, m \in \mathbb{N}$ and $\sum_{i=1}^n p_i = \sum_{i=1}^m q_i = 1$. Thereby the payoffs x_i denote pairwise distinct monetary outcomes, which occur with strictly positive probabilities $p_i > 0$ for $1 \leq i \leq n$ (we impose analogous conventions for lottery L_y 's outcomes). The decision maker evaluates monetary outcomes via a strictly increasing value function $u(\cdot)$ with $u(0) = 0$. Each state of the world $s_{ij} := (x_i, y_j)$ occurs with some objective probability $p_i q_j$.² According to standard theory, lottery L_x 's expected utility $U(\cdot)$ equals

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

²Here, it is assumed that the lotteries are independent. This assumption is for notational convenience and without loss of generality.

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 the state's objective probability $p_i q_j$ and on its salience. In particular, salience of state $s_{ij} \in S$ is determined by a symmetric, bounded and continuously differentiable salience function $\sigma(\cdot, \cdot)$ that satisfies the following three properties:

1. *Ordering*. Let $\mu = \text{sgn}(u(x_i) - u(y_j))$. 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

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

Following the smooth salience characterization proposed in Bordalo *et al.* (2012, page 1255), each state s_{ij} receives 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. We say that a state s_{ij} is the more salient the larger the salience weight assigned to it. A rational decision maker is captured by $\Delta = 1$, while the smaller Δ is, the stronger is the salience bias. An agent with $\Delta < 1$ we call 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

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

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.

The ordering property says that a state is the more salient the more the lotteries' payoffs in this state differ. In this sense ordering captures the *contrast effect*, according to which a large difference in the values assigned to the outcomes in a given state attracts a salient thinker's attention.³ Diminishing sensitivity reflects *Weber's law* of perception and implies that the salience of a state decreases if the outcomes' values uniformly increase in absolute

³If 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 increases.

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. As contrast and level effects are more intuitive and easier to understand than the properties of ordering and diminishing sensitivity we will refer to these notions whenever it is possible in the following analysis.

In Appendix B we present the closely related focusing model (KS). According to focusing, an agent’s attention directed to a given state is assessed by a focusing function (i.e., the pendant to the salience function in the preceding approach) that satisfies the contrast, but not the level effect.

3 Certainty equivalents and monotonicity

Models of choice under risk should allow to identify certainty equivalents for all lotteries 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 $L := (x_1, p_1; \dots; x_n, p_n)$ denote some lottery with $x_i \in \mathbb{R}$ for all $1 \leq i \leq n$. Outcomes are ordered such that $x_l < x_i$ for $l < i$ and probabilities p_1, \dots, p_n sum up to one.

- (a) The certainty equivalent is defined as the minimum monetary sum c which 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 $L' := (x_1, p'_1; \dots; x_n, p'_n)$ where $p'_i = p_i + \epsilon$ and $p'_l = p_l - \epsilon$ for some $i > l$ and some $\epsilon > 0$ 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 $L'' := (x''_1, p_1; \dots; x''_n, p_n)$ where $x''_l > x_l$ for some $l \in \{1, \dots, n\}$ 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 $c'' > c$.

According to Kontek (2016), in the discrete salience model certainty equivalents do not satisfy monotonicity in probabilities or may not even exist. These observations are artefacts of the simplified, discrete salience model that Kontek (2016) analyzes. Here, the objective probability of the i^{th} most salient outcome 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, monotonicity in probabilities may be violated if probability mass is shifted from a low, salient outcome to a larger, but less salient outcome which is strongly discounted.

BGS apply the simplified, discrete version of their model for analytical ease as long as the use of the underlying continuous model is not necessary. 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 well-defined 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 with certainty. Then, lottery L is (weakly) preferred over the safe option c if and only if

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

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

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

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 $p \in (0, 1)$ as $u^{-1}(\cdot)$ is strictly increasing and $U^s(L)$ is a convex combination of $u(x_1)$ and $u(x_2)$. Then,

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

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 we get

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

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

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

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 more general, discrete lottery $L := (x_1, p_1; \dots; x_n, p_n)$ with $n \geq 2$ pairwise distinct outcomes a certainty equivalent exists and is well-defined. Suppose again some choice set $\{L, c\}$ where option c gives 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 assume $x_1 < \dots < x_n$. Then a salient thinker's certainty equivalent to L is implicitly given by $c = u^{-1}(f(c))$. Analogously to the case of a binary lottery, $u^{-1} \circ f : [x_1, x_n] \rightarrow [x_1, x_n]$ 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 define a salient thinker's *risk premium* r as the difference in the value assigned to the lottery's expected value $\mathbb{E}[L]$ and the value of lottery L 's certainty equivalent c , that is $r := u(\mathbb{E}[L]) - u(c)$. 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

We investigate how salience shapes 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 explains the simultaneous demand for insurance and casino gambles. Thereby, we extend

findings by BGS to the continuous salience model. In Section 4.2, we precisely show that salient thinkers reveal a preference for skewness. We restrict our analysis to binary lotteries as these are uniquely characterized by their first three standardized central moments: expected value, variance and skewness. This allows for an unambiguous definition of skewness and a precise analysis of a salient thinker's preferences over the skewness of lotteries. We relate our findings to the growing literature on skewness preferences.

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 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 a linear value function $u(x) = x$.⁴ Then, a salient thinker chooses the safe option over the risky 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 risk averse and prefers the safe option; otherwise the agent chooses 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 downside payoff and expected value exceeds the contrast in upside payoff and expected value. Thus, we obtain

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

⁴In contrast to expected utility theory, 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).

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 more likely than the upside payoff.

On the other hand, a salient thinker might be 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. Note that ordering implies

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

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 risk seeking only if the lottery's upside x_2 occurs with a strictly lower probability than its non-negative downside x_1 . 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 $\hat{p} = \hat{p}(x_1, x_2) \in (0, 1)$ such that she prefers the safe option if and only if $p \leq \hat{p}$.*

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

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

where $\text{Var}(L) := p(1 - p)(x_2 - x_1)^2$ denotes the variance of lottery L . Other notions of skewness refer to "long and lean" tails of the risk's probability distribution. There exist several measures of skewness, which are, however, all equivalent for binary risks (Ebert, 2015, Proposition 2). Thus, only for binary risks the impact of skewness on risk attitudes 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 $p > q$ ($p < q$, $p = q$). 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 gambles, lottery games or specific investments, 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 exhibited 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.

Example 1 (Insurance). Suppose the agent has to decide whether to pay the fair insurance premium $-\mathbb{E}[L]$ 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 or 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.

Example 2 (Gambling). Suppose the agent decides whether to buy a lottery ticket at a fair price $\mathbb{E}[L]$. When participating in the lottery, she could either win a large amount (i.e., $x_2 > 0$) or nothing (i.e., $x_1 = 0$). Then, according to Proposition 2, the salient thinker buys the lottery ticket if the probability of the gain is sufficiently small. Due to diminishing sensitivity, the salient thinker prefers the gamble only if the risk is right-skewed.

Example 3 (Investments). Suppose the agent decides whether to buy an asset with positive expected value—that either pays $x_1 < 0$ or $x_2 > 0$ in the future—at its fair price $\mathbb{E}[L]$. If the probability of the gain is sufficiently high, the downside payoff x_1 stands out and the salient thinker chooses the safe option. If the probability of the loss is high, the upside payoff x_2 is salient and the salient thinker buys the asset at its fair price. This implies a tendency to buy right-skewed assets as Bordalo *et al.* (2013a) have already pointed out.⁵

4.2 Salience and skewness preferences

In line with the empirical evidence, salience theory suggests that the skewness of the risk’s probability distribution affects risk attitudes. Most field studies, however, do not

⁵While 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 derive a salient thinker’s preference for skewness formally.

precisely test for the role of skewness in risk taking as variance and skewness of typical casino gambles or lottery games are not independent, but 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 the expected value and the 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.”

Hence, 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. As for given outcomes x_1 and x_2 a change in the probability p also induces a change in the lottery’s expected value $\mathbb{E}[L]$ and its variance $Var(L)$, we cannot infer from Proposition 2 whether it is the skewness of the risk which induces the aversion toward left-skewed and the preference for right-skewed lotteries.

Lemma 1 (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*

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

For a proof of this Lemma see Ebert (2015). Using this characterization, we can assess the impact of skewness on the salient thinker’s risk attitude. As before, we assume $u(x) = x$ such that—for the binary lottery L with expected value E , variance V and skewness S —the salient thinker’s risk premium equals

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

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 over the safe option E if and only if the lottery’s risk premium is strictly negative, or, equivalently, its upside payoff is salient. We conclude:

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

Suppose the lottery’s expected value and variance are fixed. Then, equation (2) yields that 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 downside payoff and 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. By continuity of the salience function we obtain the following corollary.

Corollary 1. *For any 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 .*

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

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. Thus, 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 is smaller for σ than for $\hat{\sigma}$ if and only if the lottery is right-skewed.*

This implies that a stronger contrast effect enhances a salient thinker's aversion toward left-skewed risks and her preference for right-skewed risks. As we derive the preference for skewness from lotteries with the same expected value, the contrast effect is equivalent to ordering in this context and a salient thinker's preference for skewness is the stronger the more important ordering is relative to diminishing sensitivity.

Experimental evidence on skewness preferences. Our preceding results are in line with experimental evidence on skewness preferences. In contrast to studies with field data, laboratory experiments allow to precisely test for skewness preferences (i.e., the skewness of a lottery can be varied *certeris 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.⁶ They also show that prudence (i.e., $u'''(\cdot) < 0$) does not suffice to ex-

⁶More 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. 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 .

plain skewness preferences.⁷ 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., Br  nner *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,   stebro *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 preferences and related predictions under controlled conditions in the laboratory.

5 Puzzles on skewness preferences

In many regards, the predictions by salience theory of choice under risk meet the predictions by cumulative prospect theory (for a detailed discussion, see BGS). For instance, both theories predict that it depends on the skewness of a risk whether an agent buys insurance or prefers to gamble. The skewness of a distribution may, however, induce implausible predictions for cumulative prospect theory as shown by three more or less recent articles. On the one hand, Ebert and Strack (2015) argue that for any value function (and any reference point) there exists a right-skewed and arbitrarily small binary risk with 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 additional assumptions on the value function also arbitrarily large, unfair gambles attract agents with CPT preferences. 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 to stop as soon as she either has realized a rather small loss x_1 or a large gain x_2 . This stopping strategy with two absorbing end-points can be represented as a binary lottery which gives a small loss with a large probability, and a large gain with a small probability. According to Corollary 1, a salient thinker 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 cannot commit to a long-run stopping

⁷Prudence can explain a preference for positive skewness given a fixed expected value and variance. While subjects making prudent choices also tend to choose right-skewed lotteries in the experiment by Ebert and Wiesen (2011), prudence is not sufficient to explain the number of skewness-seeking choices.

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 naive CPT agent that uses the preceding stopping strategy will never stop gambling irrespective of her value function’s curvature.⁸ 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 the lottery’s expected value is 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 which 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 to quit her tedious job or to buy an expensive car, thereby overweighting the probability that her dream will become true.

Cumulative prospect theory’s prediction that an agent will, irrespective of her value function, play until bankruptcy has been regarded as implausible and therefore as a weakness of the model. We will show that the prediction does not necessarily hold for salient thinkers as soon as the assumption of a linear value function is dropped. Precisely, we investigate conditions under which Corollary 1 breaks down so that a salient thinker will not follow the above stopping strategy until her entire wealth is lost. In fact, if the value function is strictly concave, a salient thinker may or may not be inclined to gamble, depending on the interplay of her value function’s and her salience function’s curvature.⁹

⁸The naive 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 negative expected value and continues 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). The sophisticated agent is aware of her time-inconsistency and foresees that she will not stop to gamble. Hence she decides not to gamble in the first place.

⁹The 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.* (2016a), for instance, conduct such an estimation for the focusing model, estimating simultaneously the value and the focusing function.

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 drop our previous assumption of a linear value function 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 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 $\Delta_k := \Delta^{-\sigma(u(x_k), u(\mathbb{E}[L]))}$, $k \in \{1, 2\}$. For any given expected value $E = \mathbb{E}[L]$, substituting $p = (x_2 - E)/(x_2 - x_1)$ yields

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

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 smaller than one for any strictly 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.¹⁰ 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 as (context-independent) risk aversion becomes stronger. Compared to a linear value function, the contrast between the values assigned to the upside payoff and the expected value, respectively, is reduced. As the preference for skewness is driven by the contrast effect, 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 (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 relative importance of ordering and diminishing sensitivity changes with the level of values assigned to the outcomes. Therefore, it is not possible to make a general statement on how the value function's curvature affects a salient thinker's risk attitude (see Example 4 for an illustration).

More can be said about the properties of the salience function that facilitate risk seeking behavior. As established in Proposition 4, a salient thinker's preference for right-skewed risks is driven by the contrast effect. A salient thinker is especially prone to gamble if a large gain occurring with small probability stands in a sharp contrast to the lottery's expected value, thereby grabbing much attention. Hence a salient thinker is the more risk

¹⁰Note that $u(E) - u(x_1) \geq u(x_2) - u(E)$ for any $p \leq 1/2$ due to strict 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.

seeking with respect to sufficiently right-skewed lotteries the stronger the contrast effect is relative to the level effect. In order to verify this intuition also for a concave value function, we compare salience functions which 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 satisfies (C.1) also for salience function σ .*

In line with our previous argumentation, there does not always exist a fair-priced lottery a salient thinker opts for. If the value function is very concave and the salience function exhibits a weak contrast effect at the same time, such a lottery does not exist (i.e., condition (C.1) is never satisfied). We show this by the use of two examples for which we assume power utility $u(x) = x^\alpha$ with $\alpha \in (0, 1)$ and salience function $\sigma(x, y) = \frac{\beta(x-y)^2}{(|x|+|y|+\theta)^2}$ with $\beta, \theta > 0$. Let $\theta = 0.1$ and $\Delta = 0.7$.

Example 4 (Value function). For a linear value function, the left-hand side of (C.1) equals one and the salient thinker chooses a lottery if its upside is salient. This lottery exists by Proposition 3. Then due to continuity, condition (C.1) holds also 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 $\sigma(x, y) = \frac{(x-y)^2}{(|x|+|y|+0.1)^2}$. If the value function's curvature increases, that is, the parameter α decreases, we observe that inequality (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 $\alpha = 0.5$, Figures 1 and 2 illustrate the risk premium $r = u(\mathbb{E}[L]) - U^s(L)$ as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.

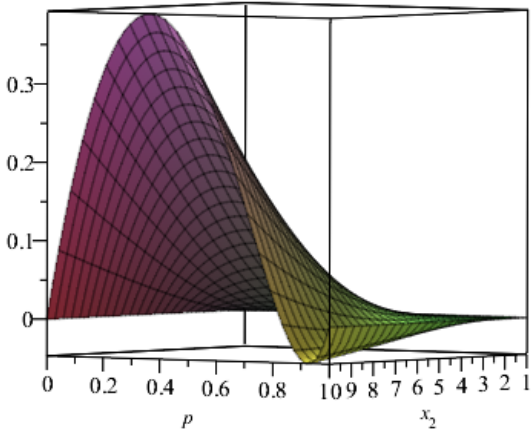


Figure 1: Risk premium for $\alpha = 0.95$.

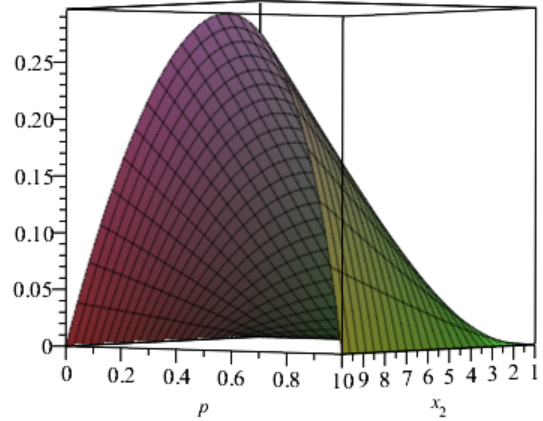


Figure 2: Risk premium for $\alpha = 0.5$.

Example 5 (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.¹¹ 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 $\beta = 10$, Figures 3 and 4 illustrate the risk premium $r = u(\mathbb{E}[L]) - U^s(L)$ as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.

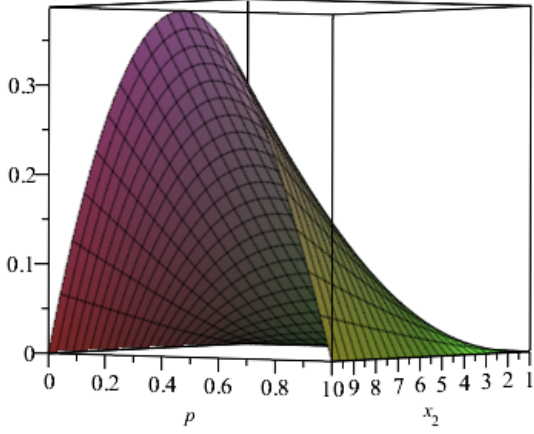


Figure 3: Risk premium for $\beta = 1$.

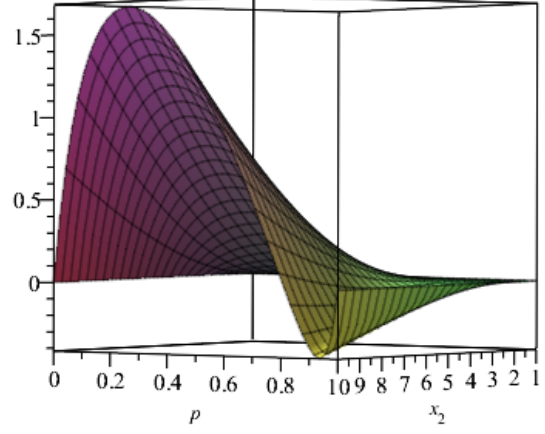


Figure 4: Risk premium for $\beta = 10$.

Note that 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) and is therefore bounded away from one, while the left-hand side of (C.1) approaches one if the variance of the lottery goes to zero. Thus, it is necessary to apply the continuous salience model in order to resolve the skewness puzzle delineated by Ebert and Strack (2015).

Dynamic salience predictions under asset integration. Suppose an agent asset integrates, that is, she evaluates a lottery's outcomes not separately, but based on the wealth levels the lottery potentially induces. If a salient thinker asset integrates, she may follow the delineated stopping strategy only until a certain wealth level is reached. We briefly discuss for which wealth levels an unfair, but attractive binary lottery might exist. Whether such a lottery exists for a particular wealth level depends on how the value function's and the salience function's curvature change in wealth. First, suppose the value function satisfies decreasing absolute risk aversion. For a suitable salience function, it may be the case that a salient thinker gambles only at high wealth levels, but stops after her wealth has sufficiently decreased. Second, if the value function satisfies increasing absolute risk aversion the opposite may arise: if the agent's wealth increases, her value function's relative curvature increases as well and it becomes more likely that no binary lottery satis-

¹¹The larger β the stronger the contrast effect is for $\sigma_\beta(x, y) = \frac{\beta(x-y)^2}{(|x|+|y|+0.1)^2}$. This, however, 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 $\sigma_\beta(x, z) - \sigma_{\tilde{\beta}}(x, z) > \sigma_\beta(y, z) - \sigma_{\tilde{\beta}}(y, z)$ if $x > y \geq z$ or $x < y \leq z$.

fies inequality (C.1). Then the salient thinker stops gambling if her wealth is sufficiently large. In contrast, for low wealth levels the curvature of the agent's value function is weak such that inequality (C.1) is more likely to hold for at least some lottery. This observation relates to the *disposition effect* (Shefrin and Statman, 1985) according to which private investors are more inclined to sell assets that have increased in value than assets which have decreased in value. Thus, salience theory might account for the disposition effect under certain circumstances.

5.2 Skewness preferences in the large

Rieger and Wang (2006) and Azevedo and Gottlieb (2012) have shown that cumulative prospect theory also yields implausible predictions for right-skewed lotteries with large absolute payoffs. Denote $\mathcal{L}(E)$ as the set of all binary lotteries with some expected value $E \in \mathbb{R}$. Azevedo and Gottlieb (2012) argue that the expected gain that can be earned by selling a lottery $L \in \mathcal{L}(E)$ to a CPT agent may be unbounded. This prediction arises from the fact that probability weighting might induce an unbounded valuation of a lottery with finite expected value (Rieger and Wang, 2006). If small probabilities are overweighted, increasing the upside payoff and reducing the corresponding probability can make a lottery more attractive. This allows a firm to realize arbitrarily large gains 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 long as we consider only lotteries with a bounded expected value. Restricting our analysis to lotteries with a bounded expected value makes sense for the following two reasons: first, Rieger and Wang (2006) argue that it is practically not feasible for a firm to offer a lottery with a very large expected value. Second, a consumer with a constrained budget is not able to pay a very large price to participate in a lottery.

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 $C < \infty$ such that $\Delta^{-\sigma(x,y)} < C$ 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. *Let $\mathcal{L}(E)$ denote the set of binary lotteries L with finite expected value $E \in \mathbb{R}$. A salient thinker's valuation for some $L \in \mathcal{L}(E)$ is bounded 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 E , choose to sell that lottery $L \in \mathcal{L}(E)$ which has the largest certainty equivalent. By Proposition 6, this is bounded such that the gain a firm can earn from selling a lottery with a fixed expected value cannot become arbitrarily large.¹²

¹²Note, however, that the profit which 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 unravelled the *contrast effect* as a plausible driver of skewness preferences. According to the contrast effect, 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. Thereby, the contrast effect induces a focus on the large, but unlikely upside gain in the case of right-skewed risks and a focus on the large potential loss in the case of left-skewed risks. As a consequence, salience theory (Bordalo *et al.*, 2012) and related approaches to stimulus-driven attention that incorporate contrast effects, such as a model of focusing (Kőszegi and Szeidl, 2013), predict skewness preferences. In contrast, a *model of relative thinking* (Bushong *et al.*, 2016) which 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.

Beside skewness preferences, in particular the salience model explains a wide range of decision biases such as the Allais paradox (Bordalo *et al.*, 2012), decoy effects (Bordalo *et al.*, 2013b) or the newsvendor problem (Dertwinkel-Kalt and Köster, 2016) 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.*, 2016b) work. In this study, we contribute to the agenda of comparing salience and cumulative prospect theory by showing that the salience-based explanation for skewness preferences has advantages over the CPT-based explanation.

First, a CPT agent exhibits a preference for skewness simply because she overweights small probabilities. This explanation is less insightful than the mechanism offered by salience theory according to which probability weights depend on the salience of the corresponding payoffs. Second, salience theory of choice under risk allows to rule out cumulative prospect theory's implausible predictions on the magnitude of skewness preferences (Rieger and Wang, 2006; Azevedo and Gottlieb, 2012; Ebert and Strack, 2015, 2016). Dynamically, a naive CPT agent will never stop to gamble until bankruptcy and will buy, but never exercise American options on assets (Ebert and Strack, 2015). Conversely, if the CPT agent is aware of her time-inconsistent behavior, she will not even acquire an option or start to gamble (Ebert and Strack, 2016). These predictions stand in stark contrast to robust empirical findings such as the disposition effect that states that options are exercised, assets are sold and gambles are quit if gains can be realized. As we have delineated, the salience-based explanation for skewness preferences does not necessarily go along with such drastic predictions. Concludingly, the salience approach offers a more compelling explanation for skewness preferences than cumulative prospect theory. In this sense, salience theory of choice under risk helps to better understand the motivation to engage in any kind of risky choices.

Appendix A: Proofs

Proof of Proposition 1. Consider some discrete lottery $L := (x_1, p_1; \dots; x_n, p_n)$ 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

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

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

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

Denote $\underline{X} := \{k \in \{1, \dots, n\} | u(x_k) \leq u(c)\}$ and $\overline{X} := \{k \in \{1, \dots, n\} | u(x_k) > u(c)\}$.

Then, we can re-write the above inequality as

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

Hence $\frac{\partial U^s(L)}{\partial u(c)} < 0$ 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 $\mathbf{x} := (x_1, \dots, x_n)$, $\mathbf{p} := (p_1, \dots, p_n)$. Then, we observe that

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

and

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

where

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

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

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

or, equivalently,

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

This inequality is always fulfilled as $\sigma_x^k \geq 0$ holds if and only if $u(c) \leq u(x_k)$. 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 $i, l \in \{1, \dots, n\}, i \neq l$. 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 eq. (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

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

which holds if and only if

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

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. \square

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

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

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 $p \leq \hat{p}$. 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 equation (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 downside (upside) payoff and 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 $\lim_{S \rightarrow \infty} x_2 = \infty > E$, we obtain

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

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 and the salient thinker chooses the risky option. This completes the proof. \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 a binary lottery L with expected value E , variance V and skewness S , denote $r(E, V, S)$ the risk premium

if salience of outcomes is assessed via σ and $\hat{r}(E, V, S)$ the risk premium if salience of outcomes is assessed via $\hat{\sigma}$. Then, it holds $r(E, V, S) > \hat{r}(E, V, S)$ if and only if

$$\frac{\sqrt{Vp(1-p)}(\Delta_1 - \Delta_2)}{p\Delta_1 + (1-p)\Delta_2} > \frac{\sqrt{Vp(1-p)}(\hat{\Delta}_1 - \hat{\Delta}_2)}{p\hat{\Delta}_1 + (1-p)\hat{\Delta}_2} \quad (\text{C.4})$$

where $\Delta_k := \Delta^{-\sigma(x_k, E)}$ and $\hat{\Delta}_k := \Delta^{-\hat{\sigma}(x_k, E)}$ for $k \in \{1, 2\}$. Re-writing (C.4) 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,

$$\frac{\Delta_1}{\Delta_2} > \frac{\hat{\Delta}_1}{\hat{\Delta}_2}.$$

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

$$\sqrt{\frac{Vp}{1-p}} = x_2 - E < E - x_1 = \sqrt{\frac{V(1-p)}{p}} \quad (\text{C.5})$$

since the contrast effect is stronger for σ than for $\hat{\sigma}$. Finally, we conclude that (C.5) holds if and only if $p < 1/2$, that is, if and only if the lottery is left-skewed (Definition 3). \square

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

$$u(x_2) - u(\mathbb{E}[L]) > u(\mathbb{E}[L]) - u(x_1). \quad (\text{C.6})$$

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.6), 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. \square

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 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

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

Here, the first inequality follows from the concavity of the value function, while the second inequality follows from using the upper bound of C for the salience weights in the numerator and the lower bound of 1 for the salience weights in the denominator. This completes the proof. \square

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 towards risky choices along the lines of salience theory, that is, the agent evaluates an option according to the underlying state space.

Model. Suppose some choice set $C := \{L_x, L_y\}$ where $L_x := (x_1, p_1; \dots, x_n, p_n)$ and $L_y := (y_1, q_1; \dots, y_m, q_m)$ 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' outcomes are pairwise 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 $p_i q_j$. 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 $p_i q_j$ 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 is 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 $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is bounded and strictly increasing.¹³

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. An agent with $\Delta < 1$ we call a *focused thinker*.

Definition 5. A *focused thinker's* decision utility $U^f(\cdot)$ for $L_x \in \{L_x, L_y\}$ is given by

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

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 $U^f(c) = U(c) = u(c)$.

¹³Relatedly, 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. Thus, the probability weight on state s_{ij} decreases in the absolute difference in values d_{ij} .

Certainty equivalents and monotonicity. Suppose the agent faces some choice set $\{L, c\}$ where $L := (x_1, p_1; \dots; x_n, p_n)$ is a lottery with $n \geq 2$ pairwise distinct outcomes 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 $x_1 < \dots < x_n$. Then a focused thinker's certainty equivalent is implicitly given by $c = u^{-1}(F(c))$. As for the salience model, we conclude that $u^{-1} \circ F : [x_1, x_n] \rightarrow [x_1, x_n]$ has at least one fixed point due to Brouwer's fixed point theorem and we obtain the following proposition.

Proposition 7 (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

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

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

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 $x_2 > x_1$ and 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 Lemma 1, a focused thinker's risk premium for the binary lottery L with expected value E , variance V and skewness S equals

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

where outcomes $x_k = x_k(E, V, S)$, $k \in \{1, 2\}$, and probability $p = p(S)$ are defined in equation (2). A focused thinker strictly prefers the lottery L 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 8 (Skewness preferences). *For given expected value E and variance V , a focused thinker strictly prefers the binary lottery over its expected value 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 $g(x_2 - E) > g(E - x_1)$. 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, $p > 1/2$. From equation (1) we conclude that a focused thinker strictly prefers the lottery over its expected value if and only if $S > 0$. \square

Hence a focused thinker seeks right-skewed but avoids left-skewed risks.¹⁴ As for 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.

Definition 6. *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 9 (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 is smaller for g than for \hat{g} if and only if the lottery is right-skewed.*

Proof. Analogous to the proof of Proposition 4. \square

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

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

¹⁴Note that, for any expected value E and variance V , a *relative thinker* (Bushong et al., 2016, see also footnote 12) prefers the binary lottery over its expected value if and only if $S < 0$. It is straightforward to show that 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.

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 $\alpha \in (0, 1)$. Further, we consider the focusing function $g(x) = 1 - \frac{1}{1+\gamma x}$ for some parameter $\gamma > 0$ and $x \in \mathbb{R}_+$. Let $\Delta = 0.7$.

Example 6 (Value function). Let $\gamma = 1$ 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 $\alpha = 0.5$, Figures 5 and 6 illustrate the risk premium $r = u(\mathbb{E}[L]) - U^f(L)$ as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.

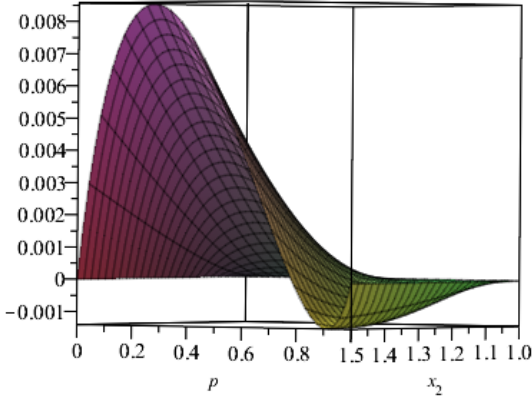


Figure 5: Risk premium for $\alpha = 0.95$.

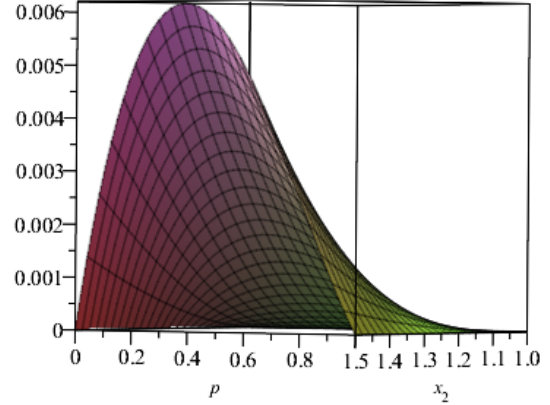


Figure 6: Risk premium for $\alpha = 0.5$.

Example 7 (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 $\gamma = 10$, Figures 7 and 8 illustrate the risk premium $r = u(\mathbb{E}[L]) - U^f(L)$ as a function of probability p and upside payoff x_2 for a given downside payoff $x_1 = 1$.

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 10. 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 $\hat{C} < \infty$

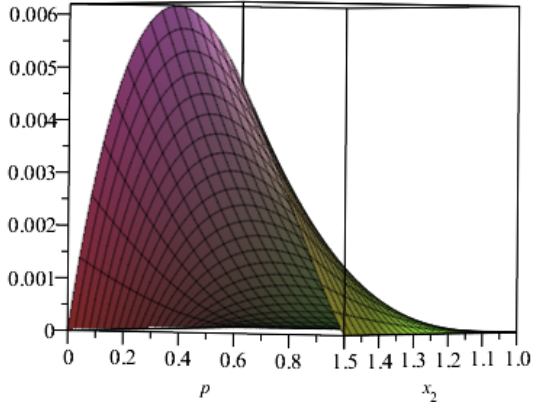


Figure 7: Risk premium for $\gamma = 1$.

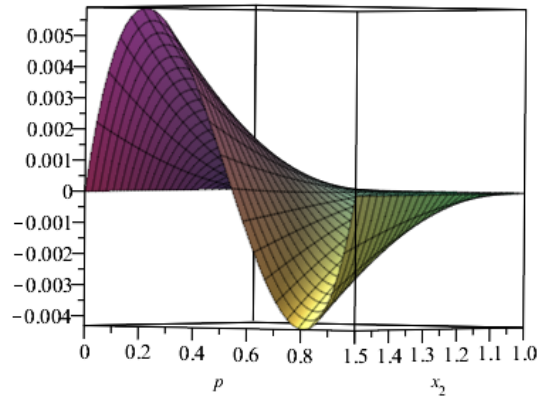


Figure 8: Risk premium for $\gamma = 10$.

such that $\Delta^{-g(x)} < \hat{C}$ for any $x \in \mathbb{R}_+$. The remainder of the proof is analogous to the proof of Proposition 6. \square

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 to identify skewness preferences.

Definition 7. Let $p \in (0, \frac{1}{2})$. Two perfectly correlated, binary lotteries $L_x := (x_1, p; x_2, 1 - p)$ and $L_y := (y_1, 1 - p; y_2, p)$ 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 11. 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—both lotteries have the same expected values—the above inequality simplifies to $\sigma(x_1, y_2) > \sigma(x_2, y_1)$. As $p < 1/2$, Lemma 1 yields

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

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

Finally, it is straightforward to see from the prove above that also a focused thinker prefers L_y over L_x 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., LEVÍNSKÝ, R. and QIU, J. (2011). Preferences for skewness: Evidence from a binary choice experiment. *European Journal of Finance*, **17** (7), 525–538.
- BUSHONG, B., RABIN, M. and SCHWARTZSTEIN, J. (2016). A model of relative thinking, 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.

- CONRAD, J., DITTMAR, R. F. and GHYSELS, E. (2013). Ex ante skewness and expected stock returns. *Journal of Finance*, **68** (1), 85–124.
- DETWINKEL-KALT, M., GERHARDT, H., RIENER, G., SCHWERTER, F. and STRANG, L. (2016a). Concentration bias in intertemporal choice, unpublished.
- , KÖHLER, K., LANGE, M. R., WENZEL, T. *et al.* (2016b). Demand shifts due to salience effects: Experimental evidence. *Journal of the European Economic Association*.
- and KÖSTER, M. (2016). Salient compromises in the newsvendor game, unpublished.
- 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, 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, 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.
- 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.
- SHEFRIN, H. and STATMAN, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*, **40** (3), 777–790.
- 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.