SALIENCE AND SKEWNESS PREFERENCES

Markus Dertwinkel-Kalt
Frankfurt School of Finance &
Management

Mats Köster
Düsseldorf Institute for Competition
Economics (DICE)

Abstract
Whether people seek or avoid risks on gambling, insurance, asset, or labor markets crucially depends on the skewness of the underlying probability distribution. In fact, people typically seek positively skewed risks and avoid negatively skewed risks. We show that salience theory of choice under risk can explain this preference for positive skewness, because unlikely, but outstanding payoffs attract attention. In contrast to alternative models, however, salience theory predicts that choices under risk not only depend on the *absolute* skewness of the available options, but also on how skewed these options appear to be *relative* to each other. We exploit this fact to derive novel, experimentally testable predictions that are unique to the salience model and that we find support for in two laboratory experiments. We thereby argue that skewness preferences—typically attributed to cumulative prospect theory—are more naturally accommodated by salience theory. (JEL: D81)

1. Introduction

Most puzzles in choice under risk can be attributed to the *skewness* (i.e., the third standardized moment) of the underlying probability distribution. First, due to

monotonicity of preferences most people like, ceteris paribus, risks with a higher expected value (i.e., a higher first moment). Second, due to intrinsic risk aversion most people dislike, ceteris paribus, risks with a higher variance (i.e., a higher second moment). But the conventional wisdom that *in general* people prefer risks with a higher expected value and/or a lower variance can be overturned by preferences over the skewness of a risk. Many individuals, for instance, overpay for insurance with low deductibles against left-skewed risks that yield a rather large loss with a small probability (e.g., Sydnor 2010; Barseghyan et al. 2013). But at the same time, these individuals often seek right-skewed risks such as casino gambles that realize a large gain with a tiny probability (e.g., Mao 1970; Kahneman and Tversky 1979; Golec and Tamarkin 1998; Garrett and Sobel 1999; Forrest, Simmons, and Chesters 2002). Also the famous Allais paradoxes are at the heart a manifestation of the tendency to avoid negatively skewed risks. The fact that people seek right-skewed and avoid left-skewed risks is often referred to as *skewness preferences*.

A compelling explanation for skewness preferences is still missing. Since Bernoulli (1738) expected utility theory (EUT) has been the predominant model of choice under risk. Given the assumption of a concave utility function,[1] however, EUT cannot explain why people behave risk-averse in some and risk-seeking in other situations. As Kahneman and Tversky (1979) have pointed out, "Choices among risky prospects exhibit several pervasive effects that are inconsistent with the basic tenets of [expected] utility theory." In response, cumulative prospect theory (CPT; Tversky and Kahneman 1992) has proposed a nonlinear probability weighting, which allows us to rationalize that people dislike some, but not all risks. Since a CPT agent overweights the probabilities of extreme events by assumption, she exhibits a preference for right- and an aversion toward left-skewed risks. Indeed, the prevalence of people who reveal a preference for positive skewness was a main reason for why CPT became the gold standard for descriptive modeling of choice under risk. Since CPT assumes that the value of an option is context-independent, however, it implies that only the *absolute* skewness of a lottery matters, but not how skewed a lottery appears to be *relative* to alternative options. We will provide experimental evidence that documents the importance of relative skewness for choice under risk, suggesting that neither EUT nor CPT offer a satisfactory explanation for skewness preferences.

As an illustration of relative skewness, let us consider the problem of whether to invest in some right-skewed asset or not. Depending on its correlation with the market (that is typically left-skewed), such a right-skewed asset appears to be more or less skewed relative to the market portfolio. If the right-skewed asset is, for instance, negatively correlated with the market, then it yields extreme, above-average returns exactly when the market performs badly, which happens to be the case only with a rather small probability. As a consequence, the asset's distribution of excess returns (relative to the market) is right-skewed, which we interpret as the asset being skewed

---

1. This assumption is necessary to explain why people avoid symmetric mean-preserving spreads over positive outcomes.

relative to the market. If the asset is, however, positively correlated with the market, then its distribution of excess returns can be left-skewed. In this case we would argue that the market is skewed relative to this asset, which suggests that a right-skewed asset is not necessarily also skewed in relative terms. Based on this idea we will propose a notion of *relative* skewness, and we will experimentally show that subjects decide on the grounds of the options' *relative* rather than *absolute* skewness. Referring to the preceding example this means that people have a stronger preference for the right-skewed asset the more negatively correlated it is with the market, and in our experiment we can rule out that this correlation-dependence of preferences is driven by reasons of diversification. As in CPT an asset's value only depends on its marginal distribution, CPT cannot account for the role of relative skewness in choice under risk.

Salience theory of choice under risk (Bordalo, Gennaioli, and Shleifer 2012) provides an intuitive account for why people like skewness, both in absolute and relative terms. Here, attention is automatically directed toward states of the world (i.e., in binary choices, toward pairs of outcomes) that stand out in the choice context whereas less salient states tend to be neglected. Similar to CPT, salience theory incorporates nonlinear probability weighting, but the distortion of a state's probability weight is determined by the contrast in the corresponding outcomes. Probabilities of outstanding states are inflated, whereas probabilities of less salient states are underweighted. In a typical lottery game, for instance, the large jackpot differs by much from the rather low price of the lottery ticket, thereby attracting a great deal of attention. As a consequence, the agent overweights the probability of winning the salient jackpot and behaves as if she was risk-seeking. In contrast, an agent typically demands insurance against unlikely, but potentially large losses. Compared to the rather small insurance premium the large loss stands out, its probability is inflated, and the agent behaves as if she was risk-averse. In summary, the salience mechanism yields both a preference for right- and an aversion toward left-skewed risks. Moreover, since in salience theory the value of an option depends on the joint distribution of all available alternatives, salience can account for the role of relative skewness in choice under risk.

Our contribution in this paper is threefold. First, we show that the salience model predicts the behavior that the literature has referred to as skewness preferences; that is, whether a salient thinker opts for a risky instead of a safe option depends in a systematic way on the skewness of the risk at hand. To be more precise, we study a salient thinker's preferences over the skewness of binary lotteries building on the methodology developed in Ebert (2015), and we find that a salient thinker is more likely to choose a binary risk over its expected value if it is ceteris paribus (i.e., for a given expected value and a given variance) skewed further to the right. By varying a lottery's skewness independent of its expected value and variance, we are able to identify a salient thinker's preferences over the skewness of a risk, which extends the analysis of salience-dependent risk attitudes provided in Bordalo et al. (2012). We also single out the channel—namely, the *contrast effect*—through which salience theory predicts skewness preferences. The contrast effect means that when comparing a risky and a safe option, a certain outcome of the risky option receives the more attention the more it differs from the safe option's payoff.

Second, when controlling for the variance of the available lotteries, we show that not only a lottery's absolute skewness is important, but it also matters how skewed this lottery is relative to the alternative option. To formalize this fact, we propose a novel measure of *relative skewness*: we say that a lottery L_x is skewed relative to a lottery L_y if and only if the difference in outcomes, $L_x - L_y$, is right-skewed. This measure of relative skewness is closely aligned to the idea of the contrast effect. By the contrast effect, a salient thinker's behavior depends on the difference in attainable outcomes in the different states of the world, and it can thereby vary with changes in a lottery's relative skewness. In particular, since the correlation of the two lotteries determines the set of feasible states of the world and thereby also the lotteries' relative skewness, the lotteries' correlation can affect a salient thinker's behavior even if it does not convey any relevant information. More specifically, we delineate under which conditions a salient thinker prefers a right- over a left-skewed binary lottery with the same expected value and the *same* variance, and we formulate novel, experimentally testable predictions that are based on our measure of relative skewness. As we argue in Appendix A, the concept of relative skewness is not only useful in the context of our experiment, but also helps us to better understand the seemingly unrelated Allais paradoxes. As it turns out, not only the common-consequence Allais paradox, but also prominent versions of the common-ratio Allais paradox (see, e.g., Kahneman and Tversky 1979; O'Donoghue and Sprenger 2018) can be explained by a preference for relative skewness.

Third, we conducted two laboratory experiments in order to test for our predictions. These experiments allow us to precisely disentangle the salience-based explanation for skewness preferences from alternative approaches such as CPT. As predicted by salience, we find that the more skewed the risky option is, the more likely it is that subjects will choose a risky option over the safe option that pays its expected value. Also in line with the salience model, this preference for positive skewness becomes stronger for lotteries with a larger expected value. In a second experiment we control for the options' variance, but vary both the lotteries' absolute skewness and their correlation structure, which enables us to identify the effect of relative skewness on choice under risk. Our experimental results confirm that relative skewness indeed plays an important role, which is consistent with salience, but not with CPT. Hence, given the empirical relevance of the lotteries' relative skewness, we argue that skewness preferences typically attributed to CPT are more naturally accommodated by the salience model.

Skewness preferences are not only relevant for insurance and gambling decisions, but they also have important implications for other economic and financial decision situations. On asset markets, for instance, investors pay a premium for positive skewness (Boyer, Mitton, and Vorkink 2010; Bali, Cakici, and Whitelaw 2011; Conrad, Dittmar, and Ghysels 2013). This may help us to understand the well-known growth puzzle (Fama and French 1992) according to which value stocks, that 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 whereas growth stocks are usually right-skewed. Along these

lines, Barberis (2013) argues that firms conducting an initial public offering (IPO) have a lower average return on their stocks than comparable firms that did not conduct an IPO, because stocks for which an IPO is conducted typically yield a right-skewed distribution of returns and are therefore overpriced. Accordingly, Green and Hwang (2012) find that the more skewed the distribution of returns is, the lower the long-term average return of an IPO stock is. Chen, Hong, and Stein (2001) even argue that managers strategically disclose information in order to create positive skewness in the distribution of stock returns. Skewness preferences also play an important role for portfolio selection (Chunhachinda et al. 1997; Prakash, Chang, and Pactwa 2003; Mitton and Vorkink 2007), and allow us to understand the prevalent use of technical analysis for asset trades, even though it is futile in light of the efficient market hypothesis (Ebert and Hilpert 2016). Finally, a preference for positive skewness also matters in labor economics. Hartog and Vijverberg (2007) and Berkhout, Hartog, and Webbink (2010) argue that workers accept a lower expected wage if the distribution of wages in a cluster (i.e., education–occupation combination) is right-skewed. Grove, Jetter, and Papps (2018) investigate factors that induce junior tennis players to pursue the risky career of a professional tennis player. Using longitudinal data they show that junior tennis players are attracted to the profession by highly right-skewed earnings distributions. In line with this evidence, Choi, Lou, and Mukherjee (2016) observe that the number of college students choosing to major in a certain field is the higher the more right-skewed the distribution of stock returns of potential employers is.

We proceed as follows. In Section 2, we present the salience model and derive first results. Section 3 introduces measures of absolute and relative skewness, which we use in Section 4 in order to delineate our main theoretical results on skewness preferences. We present two experiments on salience and skewness preferences in Section 5 and provide the corresponding experimental results in Section 6. Section 7 then discusses alternative accounts for skewness preferences. Section 8 concludes. All proofs are relegated to the Appendix.

2. A Continuous Version of Salience Theory of Choice under Risk

2.1. Model

Suppose the choice set $\mathcal{C}$ contains exactly two real-valued lotteries, L_x and L_y. The corresponding space of states of the world $S \subseteq \mathbb{R}^2$ is determined by the joint distribution of these lotteries. We denote the joint cumulative distribution function as F. If some lottery is degenerate, we call it a *safe* option. With slight abuse of notation, we refer to $L_z \in \{L_x, L_y\}$ both as a lottery and the corresponding random variable.

According to salience theory of choice under risk (Bordalo et al. 2012, henceforth: BGS), a decision-maker evaluates a lottery by assigning a subjective probability to each state of the world $s \in S$ that depends on the state's objective probability and on its *salience*. The salience of a state $s \in S$ is determined by a so-called *salience function* that is defined as follows.

DEFINITION 1 (Salience Function).   A symmetric, bounded, continuous, and almost everywhere continuously differentiable function $\sigma : \mathbb{R}^2 \to \mathbb{R}_+$ is a salience function if and only if it satisfies the following three properties:

1. Ordering. Let $\mu = \mathrm{sgn}(x - y)$. Then for any $\varepsilon, \varepsilon' \geq 0$ with $\varepsilon + \varepsilon' > 0$,

$$\sigma(x + \mu\,\varepsilon, y - \mu\,\varepsilon') > \sigma(x, y).$$

2. Diminishing sensitivity. Let $x, y \geq 0$. Then for any $\varepsilon > 0$,

$$\sigma(x + \varepsilon, y + \varepsilon) < \sigma(x, y).$$

3. Reflection. For any $x, y, x', y' \geq 0$, we have

$$\sigma(x, y) < \sigma(x', y') \quad \text{if and only if} \quad \sigma(-x, -y) < \sigma(-x', -y').$$

We say that a given state of the world $(x, y) \in S$ is the *more salient* the larger its salience value $\sigma(x, y)$ is. The ordering property implies that a state is the more salient the more the lotteries' payoffs in this state differ. Thus, ordering captures the well-known *contrast effect* (e.g., Schkade and Kahneman 1998), whereby decision makers focus their attention on those states of the world where the attainable outcomes differ a lot.[2] Diminishing sensitivity reflects *Weber's law* of perception and it implies that the salience of a state decreases if the outcomes in this state uniformly increase in absolute terms. Hence, diminishing sensitivity can be described as a *level effect* according to which a given contrast in outcomes is more salient for lower outcome levels. Finally, by the reflection property, diminishing sensitivity (with respect to zero) reflects to the negative domain. Instead of the rather abstract terminology of ordering and diminishing sensitivity, we mainly use the more intuitive notions of contrast and level effects.

Since the relative importance of the contrast and the level effect may vary with the payoff level, we need to impose more structure on the salience function in order to derive certain comparative statics with respect to a lottery's expected value (namely, Proposition 4).

DEFINITION 2 (Decreasing Level Effect).   Suppose that $x, y, z \in \mathbb{R}$ with $x + y, x + z \geq 0$. For a given salience function σ, denote

$$\varepsilon_\sigma(x, y, z) := -\frac{\frac{d}{dx}\sigma(x + y, x + z)}{\sigma(x + y, x + z)}.$$

The salience function σ satisfies a decreasing level effect if and only if both $\varepsilon_\sigma(x, y, z)$ and $\varepsilon_\sigma(-x, -y, -z)$ strictly decrease in y and z.

The decreasing level effect states that the decrease in salience due to a uniform increase in payoffs (which is due to the level effect) is smaller for states with larger outcomes in absolute terms. In general, the decreasing level effect implies that the

---

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

contrast effect is the more important relative to the level effect the larger the payoff level is. The property of a decreasing level effect is a stronger notion of what BGS define as *convexity*,[3] and it turns out that the commonly used salience functions satisfy also this stricter property.[4]

We follow BGS in assuming that the salient thinker evaluates monetary outcomes via a linear value function, $u(x) = x$. Then, using the concept of a salience function, we can define the *salience-weighted utility* of a lottery L_x that is evaluated in the choice set $\mathcal{C}$ as follows.

DEFINITION 3 (Salience-Weighted Utility).   The salience-weighted utility of a lottery L_x evaluated in the choice set $\mathcal{C} = \{L_x, L_y\}$ is given by

$$U^s(L_x \mid \mathcal{C}) = \int_{\mathbb{R}^2} x \cdot \frac{\sigma(x, y)}{\int_{\mathbb{R}^2} \sigma(s, t) \, dF(s, t)} \, dF(x, y),$$

where $\sigma : \mathbb{R}^2 \to \mathbb{R}_+$ is a salience function that is bounded away from zero.

This gives a more general version of the continuous salience model proposed by BGS. Notice that the denominator of the integrand normalizes salience-weighted probabilities so that they sum up to one. Therefore, we obtain $U^s(c \mid \mathcal{C}) = c$ for any safe option $c \in \mathbb{R}$ and any choice set $\mathcal{C}$. In words, the normalization factor ensures that a salient thinker's valuation for any safe option is undistorted, irrespective of the composition of the choice set.

EXAMPLE 1 (*Binary Lotteries*).   Suppose that $L_x = (x_1, p; x_2, 1 - p)$ and $L_y = (y_1, q; y_2, 1 - q)$. Here, the lottery L_x realizes x_1 with probability p and x_2 with probability $1 - p$, whereas the lottery L_y realizes y_1 with probability q and y_2 with probability $1 - q$. Depending on the correlation structure of the lotteries, the state space comprises two states (under perfect correlation) to four states (under independence). Denote the probability that the state $s_{ij} := (x_i, y_j)$ is realized by π_{ij}. Then, the salience-weighted utility of lottery L_x is given by

$$U^s(L_x \mid \mathcal{C}) = \sum_{s_{ij} \in S} x_i \cdot \frac{\pi_{ij} \, \sigma(x_i, y_j)}{\sum_{s_{kl} \in S} \pi_{kl} \, \sigma(x_k, y_l)},$$

and if the lottery L_y is a safe option, that is, if $L_y = (y, 1)$, the preceding formula further simplifies to

$$U^s(L_x \mid \mathcal{C}) = \frac{px_1 \sigma(x_1, y) + (1 - p)x_2 \sigma(x_2, y)}{p\sigma(x_1, y) + (1 - p)\sigma(x_2, y)}.$$

∎

---

3.   To arrive at a definition equivalent to BGS's notion of convexity we have to restrict Definition 2 to the case of $y = z$. While convexity is *not* a sufficient condition for Lemma 1 in BGS to hold, the decreasing level effect is.

4.   In Appendix C, we verify that this property holds for basically all salience functions used in the literature.

2.2. Certainty Equivalents and the Role of First-Order Stochastic Dominance

In order to meaningfully discuss risk attitudes under salience theory, we first verify that a salient thinker's certainty equivalent, and thus also her risk premium, to a lottery are well-defined.

DEFINITION 4 (Certainty Equivalent). Consider any two lotteries L and L' with finite expected values.

(a) Suppose that an agent faces some choice set $\mathcal{C} = \{L, c\}$, where c represents a safe option. We say that the safe option c is a certainty equivalent to the lottery L if and only if $U^s(L \,|\, \mathcal{C}) = c$.

(b) Let c be the unique certainty equivalent to L and c' be the unique certainty equivalent to L'. Let L' first-order stochastically dominate L. The certainty equivalent is monotonic if and only if $c' > c$.

PROPOSITION 1. *A salient thinker's certainty equivalent to any lottery L is unique and monotonic.*

As an illustration consider a binary lottery $L = (x_1, p; x_2, 1 - p)$ with outcomes $x_2 > x_1$. Since the lottery's salience-weighted utility is a convex combination of its payoffs, at least one certainty equivalent to L exists and any certainty equivalent lies between x_1 and x_2. Now suppose that the agent chooses between the lottery L and a safe option $c \in [x_1, x_2]$, and denote as $\varphi(x_i, c) := \sigma(x_i, c)(x_i - c)$ the salience-weighted difference between one of the lottery's outcomes and the safe option. We conclude that any certainty equivalent c to lottery L has to solve $p\varphi(x_1, c) + (1 - p)\varphi(x_2, c) = 0$. In addition, any increase in the safe option's payoff does not only make the safe option more attractive, but also makes the lottery less attractive: the lottery's lower payoff becomes more salient whereas its higher payoff becomes less salient. Consequently, for any fixed lottery $L = (x_1, p; x_2, 1 - p)$, the expected salience-weighted difference, $p\varphi(x_1, c) + (1 - p)\varphi(x_2, c)$, is a monotonic function of c, which proves uniqueness of the certainty equivalent. Monotonicity then follows immediately from the fact that $\partial\varphi(x_i, c)/\partial x_i > 0$ almost everywhere, which implies that, for a fixed safe option c, first-order stochastic dominance with respect to outcomes x_i translates into first-order stochastic dominance with respect to salience-weighted differences $\varphi(x_i, c)$.[5]

Given that the certainty equivalent is well-defined, we can define a salient thinker's *risk premium r* for a lottery L as the difference in the lottery's expected value $\mathbb{E}[L]$ and its certainty equivalent $c(L)$, that is, $r(L) := \mathbb{E}[L] - c(L)$. According to Proposition 1, a salient thinker's risk premium for a lottery L is well-defined and monotonic with respect to first-order stochastic dominance. Indeed, we can show that, more generally, a salient thinker's preferences over any pair of independent lotteries are monotonic with respect to first-order stochastic dominance.

---

5. Notice that in the rank-based salience model that BGS have analyzed, the certainty equivalent is not monotonic and may not even exist (Kontek 2016). The fact that the certainty equivalent is not monotonic in the rank-based model highlights that Proposition 1 is not a trivial result.

PROPOSITION 2.   *Let $\mathcal{C} = \{L_x, L_y\}$ and suppose that the lotteries L_x and L_y are stochastically independent. Then, if the lottery L_x first-order stochastically dominates the lottery L_y, we have $U^s(L_x \,|\, \mathcal{C}) > U^s(L_y \,|\, \mathcal{C})$.*

3. Measures of Absolute and Relative Skewness

In this section, we introduce concepts to measure a lottery's absolute and relative skewness. We thereby build on Ebert (2015) who has shown that for binary lotteries, various popular measures of absolute skewness coincide so that absolute skewness is unambiguously defined in this case. To assess relative skewness, we introduce a novel measure that proves useful in our analysis of skewness preferences under salience theory. We also discuss the relationship between absolute and relative skewness. Whenever we simply speak of skewness, we refer to absolute skewness.

3.1. Absolute Skewness is Unambiguously Defined for Binary Risks

Often the absolute skewness of a lottery L is associated with its third standardized central moment

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

Other notions of absolute skewness refer to "long and lean" tails of the risk's probability distribution. In general, these different notions of absolute skewness are not equivalent. For instance, Ebert (2013) delineates an example of a distribution that has a third moment of zero (which is usually interpreted as the distribution being symmetric), but that is clearly left-skewed when judged by its tails. That said, the analysis of skewness effects on choice under risk is not unambiguous for lotteries with a general distribution.

Importantly, Ebert (2015) shows that for binary risks all conventional notions of skewness are equivalent (see Proposition 2 in his paper), so that the skewness of a binary risk is well-defined.

DEFINITION 5 (Absolute Skewness).   Consider binary lotteries L_x and L_y. We say that the lottery L_x is more skewed than the lottery L_y if and only if $S(L_x) > S(L_y)$. A lottery L_x is called right-skewed (or, equivalently, positively skewed) if $S > 0$, left-skewed (or, equivalently, negatively skewed) if $S < 0$, and symmetric otherwise.

The main difficulty in identifying the effect of skewness on risk attitudes is that in general variance and skewness are not independent. As an illustration, let us consider a stylized horse race bet $L = (1/p, p; 0, 1 - p)$ for some $p \in (0, 1)$. If we fix the lottery's lower payoff to zero, then any increase in the probability p implies (i) an increase in the lottery's skewness and (ii) a decrease in the lottery's variance. More generally, as long as we keep a lottery's outcome(s) fixed, any shift in choice that is attributed to a change in skewness can be likewise attributed to a change in variance. Indeed most empirical studies on skewness effects do not properly disentangle preferences over variance from preferences over skewness. Ebert (2015) argues, for instance, that inferring skewness

preferences at the horse track from the study by Golec and Tamarkin (1998) might be misleading: increasing the skewness of a horse race bet, while holding its expected value and variance constant, does not yield a new horse race bet, but a lottery with very different properties. Conversely, Ebert concludes that "a choice between two horse-race bets is never a choice between different levels of skewness only".

In order to disentangle preferences for variance and skewness, we need to vary a lottery's skewness for a *fixed* expected value and a *fixed* variance. For that, we rely on the characterization of binary lotteries in terms of their first three moments as provided in Ebert (2015).

LEMMA 1 (Ebert's Moment Characterization of Binary Risks). *For any 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*

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

We denote the binary lottery with expected value E, variance V, and skewness S as $L(E, V, S)$.

Besides allowing us to investigate how skewness affects choice under risk, the preceding lemma also speaks to the potential richness of predictions based on binary lotteries. An immediate implication of Ebert's Moment Characterization is that any probability distribution can be approximated by a binary lottery up to its first three moments. Thus, when restricting our analysis to binary risks we still allow for three-moment approximations of arbitrary risks. Studying preferences over binary lotteries is also of intrinsic interest since many insurance and gambling applications can be modeled as a binary lottery. Finally, a good understanding of choices among binary lotteries proves useful in other domains such as in the analysis of dynamic gambling behavior (e.g., Ebert and Strack 2015, 2018; Dertwinkel-Kalt and Köster 2019).

3.2. Relative Skewness and Mao Pairs

One way to define the relative skewness of two lotteries L_x and L_y is via the third standardized central moment of their difference $\Delta_{xy} := L_x - L_y$.

DEFINITION 6 (Relative Skewness). We say that the lottery L_x is skewed relative to the lottery L_y if and only if $S(\Delta_{xy}) > 0$ holds.

As we verify in Appendix A.1, the relative skewness of binary lotteries L_x and L_y—when defined this way—depends on (i) the difference in third moments (or *skewness*), $S(L_x) - S(L_y)$, and (ii) the difference in third cross-moments (or *coskewness*), $Cos(L_y, L_x) - Cos(L_x, L_y)$. Since the coskewness of two lotteries is determined by the lotteries' joint distribution, our measure of relative skewness varies with the state space. Thereby, relative skewness captures how skewed the distribution of lottery L_x appears to be in comparison to the distribution of lottery L_y.
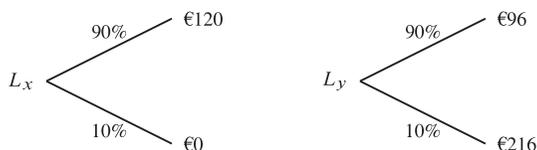
FIGURE 1. An example of a Mao pair.

At this point it might be useful to highlight some important properties of the proposed measure of relative skewness, which suggest that it is not an arbitrary choice. First, if we compare a lottery to a safe option, then our measure of relative skewness boils down to the lottery's absolute skewness. Hence, a lottery is skewed relative to a safe option if and only if it is right-skewed, which allows us to use the same skewness measure to study (1) the choice between a binary lottery and a safe option and (2) the choice between two binary lotteries. Second, when comparing two lotteries with equal variance, whether some lottery is skewed relative to the other is fully determined by the differences in third moments (i.e., skewness) and third cross-moments (i.e., coskewness). In particular, if these lotteries are stochastically independent, then a lottery L_x is skewed relative to a lottery L_y if and only if it is more skewed in absolute terms.

Next, we illustrate the concept of relative skewness using a class of binary lotteries that was introduced by Ebert and Wiesen (2011, Definition 2). This class of binary lotteries contains the lotteries used in Mao's (1970) seminal survey on skewness preferences of company managers.

DEFINITION 7 (Mao Pair). Let $S \in \mathbb{R}_+$. The lotteries $L(E, V, S)$ and $L(E, V, -S)$ denote a Mao pair.

The lotteries of a so-called *Mao pair* have the same expected value and the same variance, but differ in their direction of skewness, which makes these lotteries ideal for eliciting preferences over the skewness of a risk. Since the lotteries of a Mao pair only differ in the direction of skewness, the difference between the left-skewed lottery's lower outcome and the right-skewed lottery's lower outcome is the same as the difference between the higher outcomes of the two lotteries. As a consequence, depending on the correlation structure, the difference in the outcomes of the lotteries of a Mao pair takes at most three different values and often only two, which will prove useful in illustrating the lotteries' relative skewness (see Figure 2).

As an example of a Mao pair consider the lotteries depicted in Figure 1. Both lotteries have an expected value of $E = 108$ and a variance of $V = 1296$. The lottery L_x pays €120 with 90% probability and €0 with 10% probability; it is left-skewed. The lottery L_y pays €96 with 90% probability and €216 with 10% probability; it is right-skewed. If L_x and L_y are perfectly negatively correlated, only the states (120, 96) and (0, 216) can occur. If the lotteries are independent, all four payoff combinations (120, 96), (0, 216), (120, 216), and (0, 96) can occur. If L_x and L_y are as positively correlated as possible only the three states (120, 216), (0, 96), and (120, 96) are realized with a positive probability.
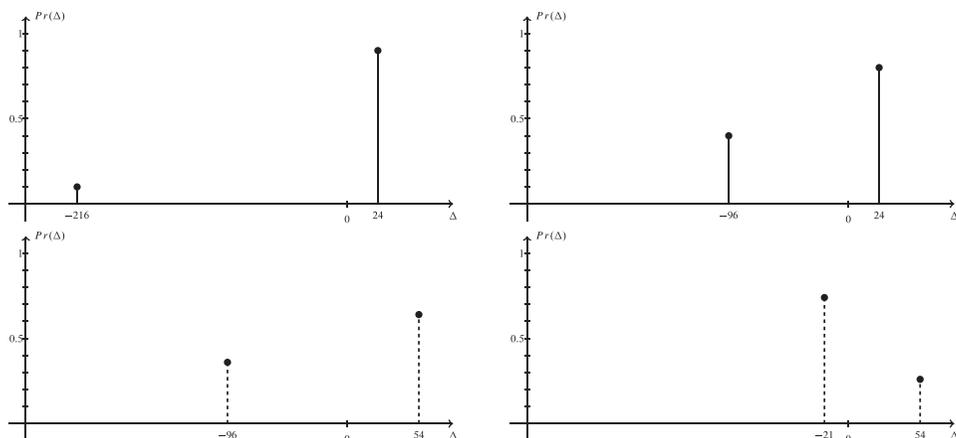
FIGURE 2. The figures in the top row depict the probability mass functions of $\Delta(V, S, \eta)$, for any Mao pair with $V = 1296$ and $S = 2.7$, under the perfectly negative correlation ($\eta = 0$, top left) and the maximal positive correlation ($\eta = 1$, top right). As both $\Delta(1296, 2.7, 0)$ and $\Delta(1296, 2.7, 1)$ are left-skewed, the lottery $L(E, 1296, 2.7)$ is also skewed relative to the lottery $L(E, 1296, -2.7)$ for both $\eta = 0$ and $\eta = 1$. The figures in the bottom row depict the probability mass functions of $\Delta(V, S, \eta)$, for any Mao pair with $V = 1296$ and $S = 0.6$, under the perfectly negative correlation ($\eta = 0$, bottom left) and the maximal positive correlation ($\eta = 1$, bottom right). As $\Delta(1296, 0.6, 0)$ is left-skewed, but $\Delta(1296, 0.6, 1)$ is right-skewed, for $\eta = 0$ the lottery $L(E, 1296, 0.6)$ is skewed relative to the lottery $L(E, 1296, -0.6)$, whereas for $\eta = 1$ the lottery $L(E, 1296, -0.6)$ is skewed relative to the lottery $L(E, 1296, 0.6)$.

TABLE 1. Joint distribution of the lotteries of a Mao pair.

Probability	ηp	$p - \eta p$	$1 - p - \eta p$	ηp
$L(E, V, -S)$	x_1	x_1	x_2	x_2
$L(E, V, S)$	y_1	y_2	y_1	y_2

In general, a Mao pair's joint distribution can be parameterized by one parameter $\eta \in [0, 1]$. Let us denote the outcomes of the left-skewed lottery $L(E, V, -S)$ as $x_1 = x_1(E, V, -S)$ and $x_2 = x_2(E, V, -S)$, respectively, and the outcomes of the right-skewed lottery $L(E, V, S)$ as $y_1 = y_1(E, V, S)$ and $y_2 = y_2(E, V, S)$, respectively. In addition, let $p = p(-S) \in (0, 1/2)$ be the probability of the left-skewed lottery's lower payoff, which is, by construction, identical to the probability of the right-skewed lottery's higher payoff. Table 1 depicts the joint distribution of a Mao pair where the parameter $\eta \in [0, 1]$ pins down the correlation structure. The correlation of the two lotteries monotonically increases in η, with $\eta = 0$ corresponding to the perfectly negative correlation and $\eta = 1$ corresponding to the maximal positive correlation.[6]

---

6. Lemma A.2 in Appendix A.1 summarizes this and further properties of Mao pairs.

Since the lotteries of a Mao pair are uniquely characterized by their first three moments and the correlation structure is identified by η, we denote a Mao pair by $M(E, V, S, \eta)$. In addition, let $\Delta(V, S, \eta) := L(E, V, -S) - L(E, V, S)$ denote the difference in outcomes for the lotteries of a Mao pair, which is by Lemma 1 independent of the lotteries' expected value.

Now we turn to the relative skewness of Mao pairs and illustrate how it depends on the absolute skewness and the correlation of the lotteries. Figure 2 depicts the distribution of $\Delta = \Delta(V, S, \eta)$, for the Mao pair introduced previously (top row) and a Mao pair with the same variance but a lower absolute skewness (bottom row), both under the perfectly negative correlation (left column) and the maximal positive correlation (right column). For starters, consider the Mao pair with a lower absolute skewness depicted in the bottom row. We observe that the distribution of the difference in outcomes is right-skewed under the perfectly negative correlation (bottom left), but left-skewed under the maximal positive correlation (bottom right). Thus, by Definition 6, the left-skewed lottery is skewed relative to the right-skewed lottery under the maximal positive correlation whereas the opposite is true under the perfectly negative correlation. For the Mao pair with a higher absolute skewness depicted in the top row, however, the right-skewed lottery is skewed relative to the left-skewed one, irrespective of the correlation structure. More formally, as we show in Lemma A.2 in the Appendix, the left-skewed lottery is skewed relative to the right-skewed lottery if and only if their correlation is sufficiently positive, that is, if and only if $\eta > (2/3)(1 + S/\sqrt{4 + S^2})$, which can only be the case for Mao pairs that are sufficiently symmetric in absolute terms, that is, only for Mao pairs with $S < (2/3)\sqrt{3} \approx 1.15$.

4. Salience and the Role of Skewness in Choice under Risk

In this section, we first show that whether a salient thinker appears to be risk seeking or risk averse depends on the skewness of the risk at hand. Put differently, for a fixed expected value and a fixed variance, a binary lottery is chosen over its expected value if and only if it is sufficiently skewed. Subsequently, we delineate that a salient thinker's choice among two binary lotteries with the same expected value and the same variance is determined by their relative skewness.

4.1. Skewness-Dependent Risk Attitudes

Suppose that an agent chooses between the binary lottery $L(E, V, S)$ and the safe option that pays the lottery's expected value; that is, the choice set is given by $\mathcal{C} = \{L(E, V, S), E\}$. The salient thinker's risk premium for the lottery $L(E, V, S)$ is given by

$$r(E, V, S) = \sqrt{\frac{V}{4 + S^2}} \cdot \left(\frac{\sigma(x_1, E) - \sigma(x_2, E)}{p\sigma(x_1, E) + (1 - p)\sigma(x_2, E)} \right), \tag{3}$$

where the outcomes $x_k = x_k(E, V, S)$, $k \in \{1, 2\}$, and the probability $p = p(S)$ are defined in equation (2). To break ties we assume that the agent chooses the lottery if and only if the risk premium is strictly negative. Then, an immediate implication of equation (3) is that a salient thinker chooses the lottery $L(E, V, S)$ over the safe option E if and only if the lottery's higher payoff is salient, which turns out to be the case if and only if the lottery is sufficiently skewed.

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

By equation (2), an increase in the lottery's skewness is equivalent to an increase in both of the lottery's payoffs and in the probability that the lower payoff is realized. Since the lottery's expected value is fixed, the difference between the lower payoff and the expected value decreases in the lottery's skewness, whereas the difference between the expected value and the higher payoff increases in the lottery's skewness. Thus, by the contrast effect, the salience of the lottery's lower (higher) payoff monotonically decreases (increases) in the lottery's skewness. We conclude that a salient thinker is the more likely to take up a binary risk the more skewed this risk is. The statement then follows from the fact that the lottery's lower (higher) payoff converges to the expected value as the lottery's skewness approaches (minus) infinity.

A useful corollary to Proposition 3 is the *fourfold pattern of risk attitudes* (Tversky and Kahneman [1992]): people are risk-averse (risk-seeking) over gambles with non-negative payoffs and a likely (unlikely) upside, and risk-seeking (risk-averse) over gambles with non-positive payoffs and a likely (unlikely) downside. Put differently, people avoid symmetric lotteries with non-negative payoffs, but seek symmetric risks with non-positive payoffs.

COROLLARY 1 (Fourfold Pattern of Risk Attitudes). *Let $\hat{S} \in \mathbb{R}$ be the cutoff value derived in Proposition* 3.

(a) *For any binary lottery with non-negative payoffs it holds that $\hat{S} > 0$.*
(b) *For any binary lottery with non-positive payoffs it holds that $\hat{S} < 0$.*

Consider a symmetric binary lottery. Due to symmetry, equation (2) gives $p = 1/2$, which implies that the difference between the low payoff and the expected value equals the difference between the high payoff and the expected value. If the lottery's payoffs are non-negative, diminishing sensitivity implies that the lower payoff is salient. Hence, a salient thinker avoids any symmetric binary risk with non-negative payoffs. Likewise, diminishing sensitivity implies that the higher payoff is salient if the lottery's payoffs are non-positive, so that a salient thinker seeks symmetric binary risks with non-positive payoffs. In any case, the salient thinker chooses a binary lottery over its expected value if it is sufficiently skewed. A lottery with non-negative payoffs has to be right-skewed

in order to be selected, while a lottery with non-positive payoffs can be attractive even if it is left-skewed.[7]

A second straightforward implication of equation (3) is that the risk premium converges to zero as the lottery's skewness becomes arbitrarily large in absolute terms. This result is driven by the fact that the salience function is bounded from above and also bounded away from zero.

COROLLARY 2. *For any expected value E and any variance V, we observe that* $lim_{S\to\pm\infty} r(E, V, S) = 0$ *holds. In particular, for any binary lottery with a fixed expected value E, the certainty equivalent*

$$c_E : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}, (V, S) \mapsto c_E(V, S) := c(L(E, V, S))$$

is bounded.

We conclude from Corollary 2 that a salient thinker's certainty equivalent to lottery $L(E, V, S)$—which corresponds to her willingness-to-pay for this lottery—is a nonmonotonic function of its skewness. In particular, for any fixed expected value, a salient thinker's certainty equivalent $c_E(V, S)$—as a function of the lottery's variance V and skewness S—is bounded from above. Although increasing a lottery's skewness boosts the salience of the lottery's higher payoff, it also lowers the probability that this higher payoff will occur. Hence, as salience effects are bounded, increasing a lottery's skewness beyond some threshold decreases a salient thinker's certainty equivalent and therefore her willingness-to-pay for a given lottery.

So far, our results are driven by *either* the contrast *or* the level effect. Next, we study the interaction of both effects, which requires us to impose more structure on the salience function. More precisely, we assume that the salience function satisfies the decreasing level effect, which allows us to derive comparative statics with respect to the lottery's expected value.

PROPOSITION 4. *If the salience function satisfies the decreasing level effect, then, for any $\varepsilon > 0$ such that $L(E, V, \widehat{S}(E, V))$ and $L(E + \varepsilon, V, \widehat{S}(E + \varepsilon, V))$ have either both positive or both negative payoffs, it holds that*

$$\widehat{S}(E + \varepsilon, V) < \widehat{S}(E, V).$$

Under the assumption of the decreasing level effect positive skewness becomes even more desirable at larger payoff levels. Since the contrast effect becomes relatively more important if the attainable outcome(s) increase, the larger contrast on a right-skewed lottery's upside becomes more salient at higher payoff levels. Consequently, a uniform increase in the payoff level—that is, an increase in the lottery's expected value from E to $E + \varepsilon$—lowers the minimum level of skewness that renders a lottery with a given variance attractive. This prediction is easily testable in the lab where we can vary the payoff level across choices (see Experiment 1 in Section 5).

---

7. BGS derive a similar result for the rank-based salience model.

4.2. Not Absolute, But Relative Skewness Shapes Choice Under Risk

Next, we try to isolate the effect of skewness on choice under risk by fixing not only the expected value but also the variance across options. Formally, we study a salient thinker's choice from the set $\mathcal{C} = \{L(E, V, S_x), L(E, V, S_y)\}$ with $S_y > S_x = -S_y$, which makes the two lotteries a Mao pair.[8] In particular, we want to understand the role that relative skewness plays in predicting a salient thinker's choice among these two lotteries. For that, recall from Section 3 that the parameter $\eta \in [0, 1]$ pins down the joint distribution of a Mao pair, and that the right-skewed lottery of this Mao pair is skewed relative to the left-skewed one if and only if $\eta < (2/3)(1 + S/\sqrt{4 + S^2})$. The following proposition shows that a salient thinker indeed exhibits a preference for relative rather than absolute skewness, which is reflected in the fact that she chooses the left-skewed lottery if and only if η exceeds a certain threshold (Part (a) of Proposition 5).

PROPOSITION 5. *For any Mao pair, there exist some $\check{\eta}(S) \in (0, 1]$ and $\check{S} \in \mathbb{R}_+$ so that the following statements hold:*

(a) *A salient thinker prefers $L(E, V, S)$ to $L(E, V, -S)$ if and only if $\eta \leq \check{\eta}(S)$.*

(b) *For $\eta = 0$, a salient thinker always prefers $L(E, V, S)$ to $L(E, V, -S)$.*

(c) *For $\eta = 1$, a salient thinker prefers $L(E, V, S)$ to $L(E, V, -S)$ if and only if $S \geq \check{S}$.*

If the right-skewed lottery of a given Mao pair is also skewed relative to the left-skewed lottery, which happens to be the case only if η is sufficiently small, then the distribution of $L(E, V, S) - L(E, V, -S)$ is right-skewed *or*, in other words, the contrast in outcomes is largest in those states of the world in which the right-skewed lottery yields the higher payoff. In particular, in the case of the perfectly negative correlation (i.e., $\eta = 0$), the high payoff of the right-skewed lottery always occurs simultaneous to the left-skewed lottery's low payoff, thereby attracting a great deal of attention and rendering the right-skewed lottery attractive to a salient thinker (Part (b) of Proposition 5). As an illustration consider the example of Figure 1: under the perfectly negative correlation the two states of the world are (120, 96) and (0, 216), where the latter is more salient according to the contrast effect, so that a salient thinker chooses the right-skewed lottery, thereby revealing a preference for relative skewness. Under a sufficiently positive correlation (i.e., for larger values of η), however, the left-skewed lottery could become skewed relative to the right-skewed lottery. In this case, there would be a larger contrast in those states of the world where the left-skewed lottery outperforms the right-skewed one, which would, due to the contrast effect, render the left-skewed lottery attractive to a salient thinker. But, as we have seen in Section 3, only if the lotteries are sufficiently symmetric the left-skewed lottery of a given Mao pair could be skewed relative to the right-skewed lottery (at least under the maximal

---

8. While Proposition 5(a) includes the case of independent Mao pairs, we analyze in Appendix A.2 the choice among two independent binary lotteries in general.

positive correlation). For the Mao pair presented in Figure 1, for instance, the right-skewed lottery is also skewed in relative terms, irrespective of the correlation structure (see the top row of Figure 2). This is different for the Mao pair with a lower absolute skewness depicted in the bottom row of Figure 2. Here, the lotteries' relative skewness changes as we move from the perfectly negative to the maximal positive correlation. Altogether, the contrast effect implies that even under the maximal positive correlation (i.e., for $\eta = 1$) a salient thinker chooses the left-skewed lottery if and only if the lotteries are sufficiently symmetric (Part (c) of Proposition 5), which again reflects a preference for relative skewness.

Whether relative skewness affects choice under risk is easily testable in a lab experiment, where we can manipulate the relative skewness of a Mao pair via the lotteries' absolute skewness and their correlation (Experiment 2 in Section 5). Such an experiment allows us to distinguish our salience-based explanation for skewness preferences from alternative approaches (e.g., CPT) that suggest that only absolute and *not* relative skewness matters (see Section 7).

5. Two Experiments on Salience and Skewness

Our analytical results give rise to novel predictions that we have tested for in two laboratory experiments. While Experiment 1 tests for skewness-dependent risk attitudes, Experiment 2 investigates the role of relative skewness in choices among the lotteries of a Mao pair.

5.1. Design and Predictions

We ensure incentive compatibility by relying only on lotteries that have non-negative payoffs. In order to improve power, we implement both experiments using a within-subjects design. Detailed instructions on both experiments can be found in the Online Appendix.

Experiment 1.   In our first experiment, subjects choose repeatedly between a binary lottery $L(E, V, S)$ and the safe option paying its expected value E, where the lottery's skewness S is gradually increased. We repeat all these choices with the only difference being that all payoffs (the safe payoff E and the lottery's payoffs) are uniformly increased by some amount $\varepsilon > 0$.

The choices between E and $L(E, V, S)$ as well as those between $E + \varepsilon$ and $L(E + \varepsilon, V, S)$ allow us to test for the prediction arising from Proposition 3, saying that for each subject there exists a certain threshold value $\widehat{S}$ so that this subject prefers a binary lottery with a fixed variance over its expected value if and only if the lottery's skewness exceeds $\widehat{S}$. Since the threshold value $\widehat{S}$ depends on the curvature of the salience function, it may vary across subjects. Thus, Proposition 3 implies that the share of subjects choosing the binary lottery $L(E, V, S)$ over its expected value E monotonically increases in the lottery's skewness S.

PREDICTION 1. *For any E and any V, the share of subjects choosing L(E, V, S) over E increases in S.*

Comparing choices from the set $\{E, L(E, V, S)\}$ with those from $\{E + \varepsilon, L(E + \varepsilon, V, S)\}$ allows us to test for Proposition 4 whereby skewness preferences become stronger at larger payoff levels. Accordingly, for each subject, the skewness level $\hat{S}$ that makes this subject indifferent between the safe option E and the lottery $L(E, V, S)$ should decrease in E. Given that subjects are heterogeneous with respect to their threshold values $\hat{S}$, we predict that the share of risk takers increases in the lottery's expected value.

PREDICTION 2. *For any V and any S, the share of subjects choosing L(E, V, S) over E increases in E.*

Experiment 2. In the second experiment, subjects choose repeatedly between two lotteries that form a Mao pair. On the one hand, for a fixed expected value E, variance V, and skewness S, we vary the lotteries' correlation structure as captured by η. On the other hand, for a fixed expected value E, variance V, and correlation structure η, we vary the lotteries' absolute skewness S. Both manipulations affect the relative skewness of the two lotteries. Given that subjects are heterogeneous with respect to their salience functions, Proposition 5 implies that the share of subjects choosing the right-skewed lottery (a) weakly decreases when we move from the perfectly negative correlation to the maximal positive correlation and (b) that this decrease is larger for more symmetric Mao pairs.

PREDICTION 3. *Consider two Mao pairs M(E, V, S', η) and M(E, V, S'', η) with S' < S''.*

(a) *For each of the Mao pairs the share of subjects choosing the right-skewed lottery is weakly larger for η = 0 (i.e., the perfectly negative correlation) than for η = 1 (i.e., the maximal positive correlation).*

(b) *The correlation effect described in* (a) *is larger for the more symmetric Mao pair M(E, V, S', η).*

As we have already discussed in the previous section, for sufficiently skewed Mao pairs, the salience model does not predict a shift in choice in response to a change in correlation. In line with the intuition provided by Figure 2, we expect to observe a shift from the right-skewed towards the left-skewed lottery when moving from the perfectly negative to the maximal positive correlation only if under the maximal positive correlation the left-skewed lottery is skewed relative to the right-skewed lottery, which is indeed the case if and only if $S < (2/3)\sqrt{3}$.

5.2. Implementation

For both experiments students were invited to our laboratory via ORSEE (Greiner 2015) and both experiments were implemented with z-Tree (Fischbacher 2007). In

TABLE 2. Lotteries used in Experiment 1.

Lottery	Exp. value	Skewness
(37.5, 80%; 0, 20%)	30	-1.5
(41.25, 64%; 10, 36%)	30	-0.6
(45, 50%; 15, 50%)	30	0
(60, 20%; 22.5, 80%)	30	1.5
(75, 10%; 25, 90%)	30	2.7
(135, 2%; 27.85, 98%)	30	6.9
(57.5, 80%; 20, 20%)	50	-1.5
(61.25, 64%; 30, 36%)	50	-0.6
(65, 50%; 35, 50%)	50	0
(80, 20%; 42.5, 80%)	50	1.5
(95, 10%; 45, 90%)	50	2.7
(155, 2%; 47.85, 98%)	50	6.9

each experiment subjects made multiple choices and, in order to warrant incentive compatibility, only one of these decisions was randomly drawn to be payoff-relevant (Azrieli, Chambers, and Healy 2018). For the payoff-relevant decision the outcome of the chosen lottery was determined by a computer simulation.

Experiment 1.   Each subject made twelve choices between a lottery and a safe option that paid the lottery's expected value. The order of all decisions was randomized at the subject level. We used experimental currency units (ECU) and a conversion ratio of 2 ECU: €1. The lotteries that we used in the experiment are listed in Table 2. All twelve lotteries have the same variance of $V = 225$. In addition, the lotteries one to six and the lotteries seven to twelve also have the same expected value and differ only with respect to their skewness.

We ran three sessions ($n = 62$) in January 2018 at the DICE experimental laboratory. The experiment lasted, on average, 25 min with average earnings of €21. Due to the lotteries' skewness and the corresponding payoff profiles, subjects earned up to €47.

Experiment 2.   Each subject made twelve choices between two lotteries that form a Mao pair. Again, we randomized the order of all decisions at the subject level. Experimental earnings were converted at a ratio of 4 ECU : €1. We used the six Mao pairs listed in Table 3, where the absolute skewness is chosen in a way that *only* for the more symmetric Mao pairs (with $S = 0.6$) the left-skewed lottery becomes skewed relative to the right-skewed lottery when moving from negative to positive correlation, but not for the more skewed Mao pairs (with $S = 2.7$). For each Mao pair, we implemented the correlation structures defined by $\eta = 0$ (i.e., perfectly negative correlation) and by $\eta = 1$ (i.e., maximal positive correlation), which gives us in total six paired choices (i.e., two choices for each Mao pair) per subject.

TABLE 3. Mao pairs used in Experiment 2.

Left-skewed lottery	Right-skewed lottery	Variance	Abs. skewness	Rel. skewness $\eta = 0$	$\eta = 1$
(120, 90%; 0, 10%)	(96, 90%; 216, 10%)	1296	±2.7	−2.7	−1.5
(135, 64%; 60, 36%)	(81, 64%; 156, 36%)	1296	±0.6	−0.6	1.0
(40, 90%; 0, 10%)	(32, 90%; 72, 10%)	144	±2.7	−2.7	−1.5
(45, 64%; 20, 36%)	(27, 64%; 52, 36%)	144	±0.6	−0.6	1.0
(80, 90%; 0, 10%)	(64, 90%; 144, 10%)	576	±2.7	−2.7	−1.5
(90, 64%; 40, 36%)	(54, 64%; 104, 36%)	576	±0.6	−0.6	1.0

The initial experiment consisted of three sessions ($n = 79$) that were conducted in February and March 2018 at the DICE experimental laboratory. On the request of a referee, we ran a replication with four sessions ($n = 113$) in November 2018. A power analysis based on the results of the initial experiment implied that a replication study requires a sample size of $n = 110$.[9] We ran the replication also at the DICE experimental laboratory, but excluded subjects that participated already in the initial experiment. We pre-registered the replication study in the AEA RCT Registry with the identifying number AEARCTR-0003545.[10] The experiment lasted 25 min on average, subjects earned up to €54, and average earnings were €18.

5.3. Related Experimental Literature

To the best of our knowledge none of Predictions 1–3 has been tested yet.

Experiment 1.   Previous experiments that studied choices between a safe and a risky option changed the variance and skewness of the risky option simultaneously and/or included more than one lottery into the choice set, both of which make it impossible to test for Predictions 1 and 2. The large experimental literature on probability weighting (e.g., Gonzalez and Wu 1999; Bruhin, Fehr-Duda, and Epper 2010) as well as the experiments on salience by BGS and Frydman and Mormann (2018) document violations of EUT that might be attributed to a preference for skewness, but do not identify the causal effect of skewness on risk attitudes. Also recent experiments that want to study skewness preferences do not offer tests of Predictions 1 and 2. Brocas et al. (2016), for instance, try to identify the determinants of skewness-dependent risk attitudes by comparing investments in an asset with normally distributed and thus symmetric returns to investments in a less profitable asset with right-skewed binary returns. Since the return distributions are chosen in a way such that the right-skewed

---

9.   Using a paired *t*-test with clustered standard errors (and a cluster size of three symmetric Mao pairs), this sample size allows us to detect an effect size of 13.5 percentage points, which is the effect size observed in our initial study for Mao pairs with $S = 0.6$ (see Table D.2 in the Online Appendix), at a significance level of 5% with a power of 80%. Based on the initial experiment we assumed an intracluster correlation of 0.48.

10.   https://www.socialscienceregistry.org/trials/3545/history/37113.AEARCTRegistry.

asset also has a lower variance than the symmetric one, however, investments in the right-skewed asset—that Brocas et al. (2016) observe for a majority of subjects—could be driven by both a lower variance or a higher skewness. The experiment by Grossman and Eckel (2015) controls for the options' variance, but does also not precisely test for decisions between a safe and a risky option. In their experiment subjects first choose from a menu of six options, one of which is safe. Subsequently, subjects are offered a second set of (nonbinary) lotteries that differ only in terms of third moments, that is, for each lottery in the first choice set there is a lottery with the same expected value and the same variance but a larger third moment in the second choice set. Subjects can then switch from their initial choice to any lottery in the second choice set and a majority of subjects do so at least once. Grossman and Eckel (2015) interpret switching towards a lottery with a larger third moment as a skewness-seeking choice. In contrast to the existing literature, our first experiment studies the choice between a safe option and a binary lottery in isolation while properly controlling for variance. This allows us to identify the causal effect of skewness on risk attitudes and provides a test of Predictions 1 and 2.

Experiment 2.   Previous experiments that investigated the role of skewness in choices between two risky options used nonbinary lotteries and/or did not specify the correlation structure. Building on seminal work by Eeckhoudt and Schlesinger (2006), a series of experiments analyzed whether preferences over risky options satisfy *prudence* (see Trautmann and van de Kuilen 2018 for a survey). Under EUT prudence is equivalent to a positive third derivative of the utility function and it implies— for a fixed expected value and a fixed variance—a preference for positive skewness. Although these studies document that a majority of subjects prefer lotteries with larger third moments, they do neither use binary lotteries (which comes with the caveat that skewness is not well-defined) nor do they manipulate the lotteries' correlation structure. More closely related is the experiment by Ebert (2015) who investigated preferences over binary lotteries with the same expected value and variance, but different levels of skewness, without describing the exact state space. In line with the experiments on prudence, Ebert (2015) finds a preference for right- over left-skewed risks, but he also observes that more absolute skewness is not attractive per se, meaning that subjects do not necessarily choose the more skewed option if both lotteries are right-skewed. As we discuss in Appendix A.2, this finding is consistent with salience theory under the assumption that subjects perceived the lotteries as being independent. Our second experiment directly builds on Ebert and Wiesen (2011) who formally introduced Mao pairs and experimentally investigated choices between these lotteries. We extend their experimental setup by comparing choices under different correlation structures.[11] This allows us to identify the causal effect of relative skewness on choice under risk and provides a test of Prediction 3.

---

11.   Relatedly, a series of papers tests salience theory against CPT using correlated versions of the Allais paradoxes (e.g., BGS; Frydman and Mormann 2018; Bruhin, Manai, and Santos-Pinto 2019).
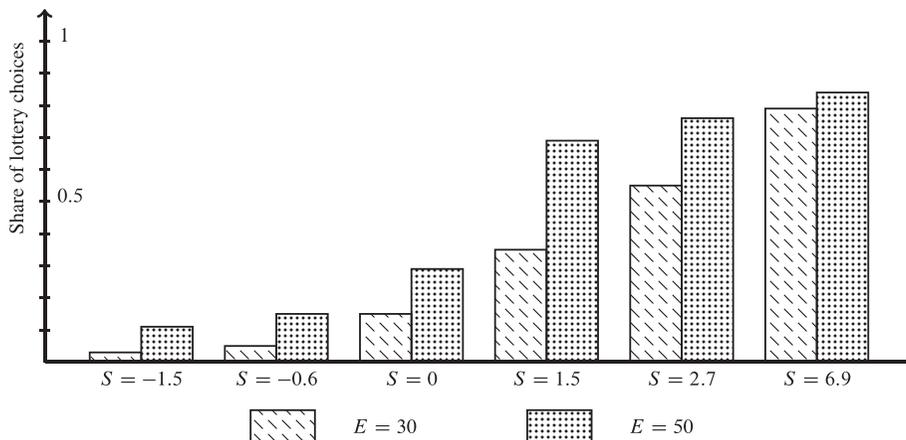
FIGURE 3. The figure depicts the share of lottery choices depending on the lottery's skewness for a low and a high expected value. The skewness values are presented in ascending order, but not in a proper scale.

6. Experimental Results

6.1. Experiment 1: Skewness-Dependent Risk Attitudes

In line with Prediction 1, the share of risk-takers strictly increases in the lottery's skewness, S, both for lotteries with an expected value of $E = 30$ (see the hatched bars in Figure 3) and for those with an expected value of $E = 50$ (see the dotted bars in Figure 3). Moreover, for all skewness levels the share of risk takers strictly increases in the lottery's expected value, E, which is in line with Prediction 2 (compare the hatched and dotted bars in Figure 3).

In order to formally test for Predictions 1 and 2, we regress a binary indicator of whether a subject chooses the lottery over its expected value on the lottery's skewness and a dummy indicating whether the lottery's expected value is high or low (see Table 4). Since each subject made twelve choices during the experiment, we cluster the standard errors at the subject level. The regression models confirm what Figure 3 suggests: the share of risk takers *significantly* increases with both a lottery's skewness and its expected value.

Returning to Figure 3, we further observe that for both payoff levels there is a particularly big shift in the share of risk takers when moving from the symmetric (i.e., $S = 0$) to the least right-skewed (i.e., $S = 1.5$) lottery. This finding is in line with the salience-based explanation of skewness preferences: the contrast on a binary lottery's upside is larger than the contrast on its downside if and only if the lottery is right-skewed, so that—due to the level effect—a binary lottery with non-negative payoffs becomes attractive to a salient thinker only if it is right-skewed. Moreover, we observe that for a low expected value the least right-skewed lottery—(60, 0.2; 22.5, 0.8)—is still chosen only by a minority, whereas the lottery that is least right-skewed

TABLE 4. Main regressions for Experiment 1.

Parameter	(1)	(2)
Constant	0.247***	0.175***
	(0.022)	(0.027)
Skewness	0.097***	0.097***
	(0.008)	(0.008)
High expected value	–	0.145***
	–	(0.025)
# Subjects	62	62
# Choices	744	744

Notes: The table presents the results of OLS regressions of a dummy indicating the choice between the lottery and the safe option (where a value of one indicates the choice of the lottery and a value of zero indicates the choice of the safe option) on the lottery's skewness and a dummy indicating whether its expected value is high or low. All standard errors are clustered at the subject level and provided in parenthesis. ***Significant at 1%.

in the set of lotteries with a larger expected value—(80, 0.2; 42.5, 0.8)—is selected by a large majority. The salience model can also explain this big shift in the share of risk takers: due to the level effect the lower payoff of lottery (60, 0.2; 22.5, 0.8) is slightly more salient than the higher payoff of this lottery although there is a larger contrast on the lottery's upside, whereas the decreasing level suggests that the upside payoff becomes salient for the lottery (80, 0.2; 42.5, 0.8). A detailed overview of the results is provided in Table D.1 in the Online Appendix.

Finally, we scrutinize whether the choice patterns predicted by Propositions 3 and 4 also hold at the individual level. We observe that 63% of all subjects (39 out of 62 subjects) have a unique switching point consistent with Proposition 3, both for the set of lotteries with a low and a high expected value, which is a much larger share than we would expect under random choice. Among the remaining 23 subjects there are 18 subjects who have a unique switching point in line with Proposition 3 for exactly one set of lotteries (i.e., either for the lotteries with a low or for those with a high expected value), and only one subject has a unique switching point in the opposite direction (i.e., choosing the lottery if and only if it is sufficiently left-skewed) for both sets of lotteries.[12] Moreover, we observe that 89% of those subjects who reveal a unique switching point for both sets of lotteries also chose the lottery weakly more often (and 51% of these subjects chose it strictly more often) in the case of a high expected value, which is consistent with Proposition 4.

In summary, the results of our first experiment confirm the importance of *absolute* skewness for risk attitudes as well as our salience-based explanation of skewness preferences.

---

12. A more formal treatment of skewness effects at the individual level is provided in Figure D.5 in the Online Appendix, where we plot the point estimates for the coefficient on the lottery's skewness in individual-level versions of Regression Model (1) in Table 4. Although in line with Proposition 3 the majority of point estimates is positive, the confidence intervals have to be interpreted with caution due to the small number of observations per subject. The results remain basically the same when using Regression Model (2) from Table 4.
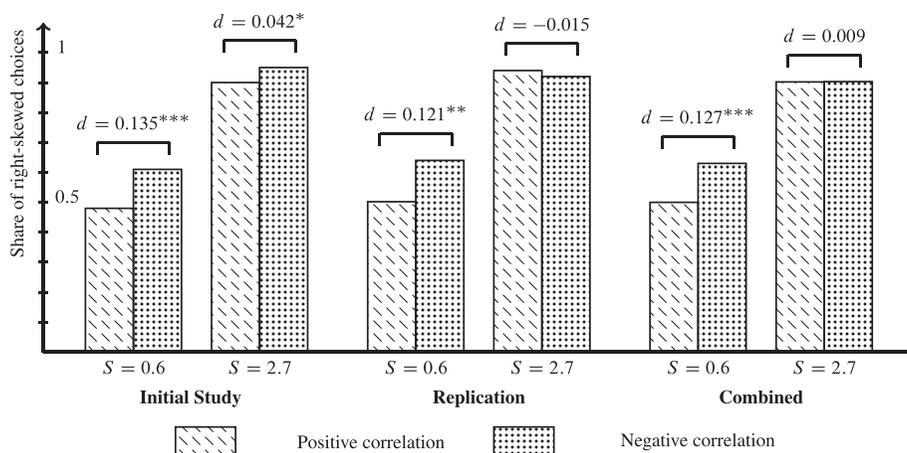
FIGURE 4. The figure illustrates the share of choices of the right-skewed lottery under positive and negative correlation. We present results separately for the initial study and the replication as well as the combined results for both studies. We further report results of paired t-tests with standard errors being clustered at the subject level (see Table 5 for the full regression). Significance level: * Significant at 10%; ** significant at 5%; *** significant at 1%.

6.2. Experiment 2: A Preference for Relative Skewness

First, when pooling all choices for the more symmetric Mao pairs (i.e., those with $S = 0.6$), we observe that the right-skewed lottery is chosen much more often under the perfectly negative correlation than under the maximal positive correlation. Second, when pooling all choices for the more skewed Mao pairs (i.e., those with $S = 2.7$), we do not find much of a difference in the frequency with which the right-skewed lottery is chosen under the perfectly negative and the maximal positive correlation, respectively. Both findings are consistent with Prediction 3 and, as illustrated in Figure 4, do not only hold for our initial study, but are successfully replicated in a properly powered replication. Table D.2 in the Online Appendix provides a detailed overview of the results. Moreover, using the combined data for the initial study and the replication,[13] we show in the Online Appendix that the results are robust across the different Mao pairs (see Figure D.6 in the Online Appendix).

Based on our power calculation—for which we used a paired t-test with standard errors being clustered at the subject level—we construct the dependent variable for a linear regression model as follows. First, we create for each choice (i.e., twelve per subject) a dummy variable that takes a value of one if the right-skewed lottery is chosen and a value of zero otherwise. Second, using these dummies, we construct

---

13. Since we determined the necessary sample size for the replication in order to detect the same effect size for the symmetric Mao pairs as observed in our initial study under the assumption that each subject makes three paired choices, we do not have enough power to test our hypothesis for each symmetric Mao pair separately when using only the data from the replication.

TABLE 5. Main regressions for Experiment 2.

Parameter	Initial study	Replication	Combined
Constant	0.135***	0.121**	0.127***
	(0.046)	(0.047)	(0.034)
Skewed	− 0.093*	− 0.136**	− 0.118***
	(0.054)	(0.053)	(0.038)
# Subjects	79	113	192
# Paired choices	474	678	1,152

Notes: The table presents the results of OLS regressions of the shift in choice (i.e., the difference in choices under the perfectly negative correlation and the maximal positive correlation) on a constant (which corresponds to the mean shift in choice for the more symmetric Mao pairs and therefore tests for Part (a) of Prediction 3) and a dummy indicating skewed Mao pairs (which gives the difference in means between the more symmetric and the more skewed Mao pairs and therefore tests for Part (b) of Prediction 3). The first column uses only data from the initial study, the second column uses only data from the replication, and the third column combines both datasets. All standard errors are clustered at the subject level and provided in parentheses. *Significant at 10%; **significant at 5%; ***significant at 1%.

for each paired-choice (i.e., six per subject) a variable that indicates the *shift in choice* due to a change in correlation and can take three values: a value of *one* if a subject switches from the right-skewed lottery under the perfectly negative correlation to the left-skewed lottery under the maximal positive correlation, a value of *minus one* if a subject switches in the opposite direction, and a value of *zero* if a subject does not change her choice. Notice that the mean shift in choice corresponds to the difference in the shares of subjects switching in either direction. Then, Prediction 3(a) implies that the average shift in choice is positive (i.e., there are more subjects who switch from the right-skewed lottery under the perfectly negative correlation to the left-skewed lottery under the maximal positive correlation than subjects who switch in the opposite direction), and Prediction 3(b) suggests that the average shift in choice is strictly larger for the more symmetric Mao pairs (with $S = 0.6$).

Now in order to formally test for Predictions 3(a) and (b), we regress the shift in choice as constructed above on a constant—which indicates the mean shift in choice for symmetric Mao pairs and therefore tests for Prediction 3(a)—and a dummy *Skewed* indicating whether a given Mao pair is more symmetric (i.e., *Skewed* =0 if $S = 0.6$) or more skewed (i.e., *Skewed* =1 if $S = 2.7$)—which yields the difference in means between symmetric and skewed Mao pairs and therefore tests for Prediction 3(b). This regression (as presented in Table 5) computes paired *t*-tests of the average shift in choice being different from zero, and it also allows us to account for the fact that each subject makes multiple choices by clustering the standard errors at the subject level. In line with Prediction 3(a), we observe that for the symmetric Mao pairs the share of choices of the right-skewed lottery *significantly* decreases by around 12–13 percentage points when moving from the perfectly negative to the maximal positive correlation (i.e., the constant takes a value between 0.121 and 0.135, and it is always significantly different from zero). And, in line with Prediction 3(b), we also find that the shift in choice is *significantly* smaller for the more skewed Mao pairs (i.e., the coefficient
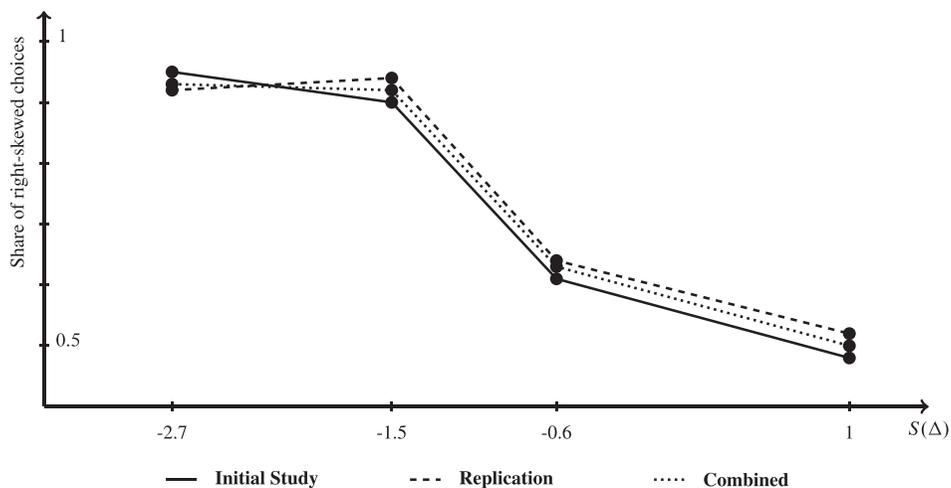
FIGURE 5. The figure illustrates the share of choices of the right-skewed lottery as a function of the left-skewed lottery's relative skewness $S(\Delta)$, where $\Delta = L(E, V, -S) - L(E, V, S)$, separately for the initial study and the replication as well as for the combined data. We observe that the share of choices of the right-skewed lottery declines in the left-skewed lottery's relative skewness.

of the dummy takes a value between -0.093 and -0.136, and it is always (weakly) significantly different from zero). As illustrated in Figure 4, for the more skewed Mao pairs there is at most a weakly significant shift in choice, that vanishes once we take all the data into account.

The preceding results provide a test of the qualitative salience predictions derived in this paper, but the data reveals further interesting patterns that are consistent with salience theory and a preference for relative skewness. While for the more skewed Mao pairs the right-skewed lottery is chosen in more than 90% of the cases, both under the perfectly negative and the maximal positive correlation, for the more symmetric Mao pairs even under the perfectly negative correlation the right-skewed lottery is chosen in only around 60% of the cases (see Table D.2 in the Online Appendix). Since the difference in salience-weighted utility between the right-skewed and the left-skewed lottery under the perfectly negative correlation is much smaller for the more symmetric Mao pairs, this finding is easily reconciled with our salience model given an appropriate assumption on the decision noise. A more intuitive way to grasp this result, however, is in terms of the lotteries' relative skewness. Recall that the relative skewness of two lotteries depends on both their absolute skewness and their correlation structure. As illustrated in Table 3, the left-skewed lottery is *less* skewed relative to the right-skewed lottery for the more skewed Mao pairs under the maximal positive correlation than it is for the more symmetric Mao pairs under the perfectly negative correlation. Along these lines, Figure 5 shows that the right-skewed lottery is chosen less often the more skewed the left-skewed lottery is in relative terms. Also a linear regression confirms that the average probability of choosing the right-skewed lottery significantly decreases with a slope of around 12.6–14 percentage points in the

relative skewness of the left-skewed lottery (see Table D.3 in the Online Appendix). Altogether, these results reveal a strong preference for relative skewness.

Finally, using the combined data from the initial study and the replication, we analyze whether the choice patterns predicted by Proposition 5 also hold at the individual level. First, we observe that 60% of all subjects (i.e., 115 out of 192 subjects) either switched from the right-skewed lottery under the perfectly negative correlation toward the left-skewed lottery under the maximal positive correlation or did not change their behavior for *all* six Mao pairs. This behavior is fully in line with Proposition 5 and a share of 60% is much larger than the share predicted by random choice. In addition, we find that 77% of all subjects (i.e., 147 out of 192 subjects) switched weakly more often in the direction predicted by salience theory. In order to study preferences for relative skewness more directly we ran individual-level regressions of the dummy indicating the choice between the right-skewed and the left-skewed lottery of a given Mao pair on the left-skewed lottery's relative skewness and Figure D.7 in the Online Appendix depicts the point estimates. Although in line with Proposition 5 the majority of point estimates is negative, the confidence intervals have to be interpreted with caution due to the small number of observations per subject.

In summary, the results of our second experiment confirm the importance of *relative* skewness for choice under risk and support our salience-based explanation of skewness preferences. Moreover, if some subjects noticed that each decision problem occurred twice in the experiment (with the only difference being that payoff-irrelevant correlations are modified) and therefore tried to choose in a consistent manner, we might actually underestimate the effect of relative skewness on choices. In this sense, due to the within-subjects design, our estimates of the average correlation effect on choice might constitute a lower bound.

7. Alternative Explanations of Skewness Preferences

7.1. Expected Utility Theory

In order to explain the fact that most people are risk-averse with respect to symmetric mean-preserving spreads over positive outcomes, EUT needs to assume that the utility function is strictly concave (Bernoulli 1738). Under this assumption, however, EUT cannot account for skewness-dependent risk attitudes as elicited in Experiment 1. More specifically, it cannot explain why—depending on the skewness of a risk—people like variance in some, but dislike variance in other decision situations. Although EUT can in principle explain why people prefer right-skewed over left-skewed lotteries with the same expected value and variance (e.g., Menezes, Geiss, and Tressler 1980; Ebert 2015), it cannot account for the correlation effects that we detected in Experiment 2. According to EUT a subject's choice between two lotteries should be independent of the correlation structure. In summary, EUT may explain a preference for absolute skewness when the variance is fixed across options, but it cannot account either for the fact that a lottery's skewness affects risk attitudes or for the role of relative skewness in choice under risk.

7.2. Cumulative Prospect Theory

In response to the weaknesses of EUT as a descriptive model of choice under risk, Kahneman and Tversky proposed *prospect theory* (Kahneman and Tversky 1979) and later developed it into *cumulative prospect theory* (Tversky and Kahneman 1992). CPT fundamentally deviates from EUT in two directions by assuming that people (1) are loss-averse with respect to some reference point and (2) weight probabilities according to a nonlinear function. In order to guarantee that preferences over lotteries satisfy first-order stochastic dominance, CPT assumes that outcomes are ranked before probability weights are computed, which in turn implies that the size of the outcomes affects probability weighting in an ordinal way.[14]

CPT can explain why decision makers seek positively skewed and avoid negatively skewed risks. More specifically, given an S-shaped value function with a sufficiently positive third derivative and an inverse S-shaped weighting function that overweights not only small but also moderate probabilities, CPT can account for both results of Experiment 1. The nonlinear probability weighting function gives the effect of a lottery's skewness on risk attitudes whereas the curvature of the value function yields the effect of the expected value on risk attitudes.

In contrast to the salience model, however, in CPT the value of a given lottery is independent of the choice context in the sense that it only depends on the lottery's marginal probability distribution. As a consequence, CPT cannot explain the role that the correlation structure plays for the choice between two lotteries and, therefore, it cannot account for the results on relative skewness in Experiment 2. In other words, CPT predicts that only the lotteries' absolute skewness and *not* their relative skewness matters for choice under risk.

Finally, there are further theoretical arguments that favor the salience-based over the CPT-based explanation of skewness preferences. Unlike the salience model, CPT inherently produces puzzling predictions on the strength of skewness effects. Rieger and Wang (2006) and Azevedo and Gottlieb (2012) show that, according to CPT, a firm selling sufficiently right-skewed lotteries with a fixed expected value can earn arbitrarily large profit margins in expectation. This result is based on the fact that increasing a lottery's upside payoff and simultaneously reducing the corresponding probability increases a CPT agent's willingness-to-pay for this lottery, namely due to the overweighting of small probabilities. This prediction arises for "virtually all functional forms that have been proposed in the literature" (Azevedo and Gottlieb 2012, p. 1294). Since, by Corollary 2, a salient thinker's certainty equivalent to any lottery with a fixed expected value is bounded, this puzzling prediction does not arise in the salience model.[15] Also unrealistic predictions of CPT on dynamic gambling and investment behavior (Ebert and Strack 2015, 2018) can be ruled out in the salience model (Dertwinkel-Kalt and Köster 2019).

---

14. This rank-dependence was first proposed by Quiggin (1982).

15. It is straightforward to show that the certainty equivalent to a lottery with an expected value E is bounded by a function that is affine in E.

7.3. Regret

Also in regret theory (Bell 1982; Loomes and Sugden 1982, 1987; Diecidue and Somasundaram 2017; Gollier forthcoming) context matters for valuations. Given a binary choice set, regret theory assumes that the unchosen alternative directly affects the decision maker's utility via a regret/rejoice term—added to a standard utility function—which captures the regret (rejoice) a decision maker anticipates to feel when the unchosen option yields a higher (lower) payoff in the realized state of the world. The disutility from regret (the utility from rejoice) is defined to be monotonically increasing in the difference in outcomes in a given state, so that regret models share the mathematical properties of the contrast effect. Lanzani (2018) shows that, for binary choices, salience theory is a special case of generalized regret theory by Loomes and Sugden (1987), which implies that based on choice data alone salience and regret are indistinguishable in our experiment. The psychology behind regret theory is, however, very different from the psychology behind salience theory. To affect behavior, regret has to be anticipated, which necessitates that the subject expects to receive information on the counterfactual outcome, as summarized in a review by Zeelenberg (1999, p. 103): "When [post-decisional] feedback is present people anticipate possible regret, but when it is absent regret does not play a significant role in the decision process." Thus, additional information on the choice process might allow us to disentangle salience and regret theory also for binary choices.

Along these lines, even though we cannot formally exclude it, we do not regard regret theory as a plausible driver of our experimental results. Consider, for instance, Experiment 1 where subjects had to choose between a safe and a risky option. Here, conditional on choosing the safe option, subjects should not expect to learn the counterfactual outcome. In fact, if a subject chooses the safe option, the lottery will not even be played out, so that there is *no* counterfactual outcome. When choosing the lottery, in contrast, the subject knows the counterfactual outcome; namely, the value of the safe option.[16] As argued by Zeelenberg (1999), regret-averse subjects should therefore choose the safe option in order to avoid any feelings of regret. In this sense, models based solely on regret aversion (e.g., Strack and Viefers 2017) cannot plausibly account for our results. Regret models that also include rejoice can, in principle, explain why subjects reveal skewness preferences in our first experiment. But despite the massive literature on regret, we are not aware of any psychological evidence substantiating rejoice.

In other experiments where salience effects have been documented, observing choices within a given context suffices to distinguish between salience and regret theory, because more than two options are presented and/or the consequences of choosing a certain option are deterministic. For instance, regret theory cannot account

---

16.   In the instructions we write: "If you have chosen [the safe option] in this task you will receive the according sum. If you have chosen [the lottery] your payoff will be determined through the simulation of the turn of a wheel of fortune. Your payoff will be paid in cash at the end of the experiment".

for various effects that decoys have. State-wise dominated decoys yield the well-established effect of asymmetric dominance (as introduced by Huber, Payne, and Puto 1982), which salience theory, but not regret theory can account for (see Dertwinkel-Kalt and Köster 2017). Also the effect of nonavailable decoys, so-called phantom decoys (e.g., Soltani, De Martino, and Camerer 2012) can be explained by salience, but not by regret (see Frydman and Mormann 2018). Moreover, regret theory makes predictions only for choices involving risk, that is, regret cannot give an alternative explanation for salience effects in risk-free domains as delineated in Bordalo et al. (2013b) and experimentally supported by Dertwinkel-Kalt et al. (2017).

7.4. Context-Dependent Attention Models

We are not the first to draw the connection between salience and skewness preferences. Bordalo et al. (2013a) have already pointed out that the salience model can explain why individuals like right-skewed and dislike left-skewed assets, but they have not disentangled a salient thinker's preferences for variance and for skewness, respectively.

Not only the salience model, but also other approaches that model context-sensitive behavior and that build on the contrast effect can account for many of our findings. For instance, our results on skewness effects carry over to the closely related *focusing model* (Kőszegi and Szeidl 2013) if it is applied to choice under risk. In this case, a focusing function—that is, the pendant to the salience function satisfying only the contrast and not the level effect—determines which states an agent's attention is directed to. As the level effect can be included into the value function (Kőszegi and Szeidl 2013, Section III.D), the focusing model can be closely aligned to the salience model in which case it shares salience theory's predictions for choice under risk. Namely, this version of the focusing model is mathematically equivalent to the generalized regret theory (Loomes and Sugden 1987).

Since our results are driven by the contrast effect, the model of relative thinking by Bushong, Rabin, and Schwartzstein (2017), which builds on the setup by Kőszegi and Szeidl (2013), but assumes a reverse contrast effect (i.e., the attention a state attracts decreases in the range of attainable payoffs in this state), cannot account for skewness preferences.

7.5. Optimal Expectations

According to the model on optimal expectations proposed by Brunnermeier and Parker (2005), an agent receives utility not only from her actions, but also from her beliefs over the likelihood of favorable future outcomes. Therefore, an agent intentionally inflates the "perceived likelihood" of upside events in order to enhance the pleasure from expecting these events. As a consequence, a model of optimal expectations predicts an excessive demand for right-skewed lotteries. But this model yields weaker predictions on skewness preferences than our salience-based approach (see, e.g., Proposition 2 in Brunnermeier and Parker 2005). First, Brunnermeier and Parker explain a preference for sufficiently right-skewed risks, but they do not obtain precise

predictions on the demand for less skewed or left-skewed assets. Second, utility from pleasant expectations can be obtained only *before* an event is realized. Thus, it seems plausible that optimal expectations matter only when there is a considerable amount of time between an investment decision and the event realization. Our salience-based approach instead explains skewness preferences irrespective of whether the realization of outcomes is delayed or not. More importantly, also a model of optimal expectations implies that only the absolute skewness of a lottery matters, but not how skewed it is relative to alternative options. Thus, Brunnermeier and Parker (2005) cannot account for the results of our second experiment.

8. Conclusion

Preferences over the skewness of the underlying probability distribution are a robust observation not only in humans, but also in animals (Strait and Hayden 2013; Genest, Stauffer, and Schultz 2016) . Choices on, for example, gambling, insurance, asset, and labor markets are crucially affected by skewness preferences. As a consequence, it is important to understand the mechanism driving skewness effects. In this paper, we have identified the contrast effect as a plausible driver of skewness preferences. Accordingly, when comparing a risky and a safe option, an outcome of the risky option attracts the more attention the more it differs from the safe option's payoff. Thereby, the contrast effect induces a focus on the large, but unlikely upside of right-skewed risks, and a focus on the large potential loss in the case of left-skewed risks. Alongside our theoretical results, we offer a novel set of experimental predictions that we have found support for in two laboratory experiments.

Our two experiments show that (1) when choosing between a risky and safe option subjects exhibit a preference for positive skewness that becomes more pronounced as the payoff level increases and that (2) the choice between two lotteries crucially depends on their correlation structure, even though this correlation is irrelevant for subjects' earnings. As Bordalo et al. (2013a) have already pointed out, the results of our first experiment might allow us to better understand, for instance, the countercyclical relationship between the aggregated stock market returns and the current economic situation. Proposition 4 suggests that we should observe more risk-seeking behavior at higher payoff levels, as for higher payoffs a salient thinker focuses even more on the large upside of a right-skewed risk and less on its downside. We confirm this prediction in our first experiment, which can also explain why in the aggregate stocks are often overvalued in boom times, but undervalued in the times of a bust (e.g., Campbell and Shiller 1988; Guiso, Sapienza, and Zingales 2018). In our second experiment we manipulated the correlation structure of the available lotteries, as this allows us to test salience theory against alternative models of skewness preferences where valuations are context-independent, such as cumulative prospect theory. Since in reality the correlation structure often carries important information and therefore *should* affect behavior, however, it is harder to come up with practical applications of our second experiment. Interestingly, although we show that making the correlation

structure explicit affects behavior even if it should not, a recently growing literature has detected that in (rather complex) choice situations where correlation indeed matters subjects tend to neglect it, thereby forming incorrect conditional probabilities when making inferences (e.g., Levy and Razin 2015; Enke and Zimmermann 2019). One conclusion that might be drawn from our experimental results is that helping people to overcome *correlation neglect* and to thereby learn the objective probabilities of the states of the world does not necessarily improve people's decisions as salience effects can distort choices even (more) if objective probabilities are known.

A potential limitation of our study is that we restrict our analysis of skewness preferences to the class of binary risks. But despite the fact that skewness is not well-defined for lotteries with a general distribution, the basic insights derived in this paper should carry over to a broader class of distribution functions. Consider for instance a symmetric distribution with continuous and bounded support on the positive real numbers (e.g., a truncated normal distribution). When choosing between this symmetric risk and a safe option paying its expected value, a salient thinker goes for the safe option due to the level effect. Now extend the support of the distribution's right tail and shift some probability mass there, which skews the distribution to the right. When compared to the risk's expected value, outcomes in the right tail attract a salient thinker's attention (due to the contrast effect) and render the risk attractive. Conversely, skewing such a symmetric distribution to the left by extending the support of its left tail makes the risk less attractive to the salient thinker, due to the outcomes in the left tail attracting an overproportionate amount of attention. Although the intuition why salience can explain skewness preferences, thus, appears to be quite robust, general results as those derived in this paper cannot be expected to hold, since for continuous distributions the salience predictions hinge on the precise definition of skewness and on the exact curvature of the salience function over the entire realm.

To sum up, our paper adds to the theoretical (Bordalo et al. 2012, 2013a) and experimental literature on salience effects in choice under risk (e.g., Dertwinkel-Kalt and Köster 2017; Frydman and Mormann 2018). Using both theoretical arguments and experimental data, we argue that skewness preferences typically attributed to cumulative prospect theory are more naturally accommodated by salience theory.

Appendix A: Relative Skewness and Choice Under Risk

A.1. Relative Skewness of Binary Lotteries and Mao Pairs

In this section, we derive some properties of our measure of relative skewness. Before we can state these properties, however, we have to introduce the concept of *coskewness*.

DEFINITION A.1 (Coskewness). The coskewness of a lottery L_x relative to a lottery L_y is given by

$$Cos(L_x, L_y) := \frac{\mathbb{E}[(L_x - \mathbb{E}[L_x])(L_y - \mathbb{E}[L_y])^2]}{\sqrt{Var(L_x)}Var(L_y)}.$$

The coskewness of two lotteries refers to their third (nontrivial) cross-moments and it is often used in the asset-pricing literature. Using the concept of coskewness, we can make the following statements on the relative skewness of two binary lotteries L_x and L_y. Recall that $\Delta_{xy} = L_x - L_y$.

LEMMA A.1. *Consider any two binary lotteries* $L_x = L(E_x, V_x, S_x)$ *and* $L_y = L(E_y, V_y, S_y)$. *Then, it holds that*

$$S(\Delta_{xy}) = \frac{S_x \sqrt{V_x}^3 - S_y \sqrt{V_y}^3 + 3Cos(L_x, L_y)\sqrt{V_x V_y} - 3Cos(L_y, L_x)V_x\sqrt{V_y}}{\sqrt{V_{xy}}^3},$$

where V_{xy} *gives the variance of* Δ_{xy}. *In addition, the following statements hold:*

(a) *If both lotteries have the same variance, then we obtain*

$$S(\Delta_{xy}) = \left(\frac{V_x}{V_{xy}}\right)^{3/2} [S_x - S_y + 3Cos(L_x, L_y) - 3Cos(L_y, L_x)].$$

(b) *If the two lotteries are stochastically independent, then we obtain*

$$S(\Delta_{xy}) = \frac{S_x \sqrt{V_x}^3 - S_y \sqrt{V_y}^3}{\sqrt{V_x + V_y}^3},$$

which simplifies to $S(\Delta_{xy}) = (S_x - S_y)/(2\sqrt{2})$ *in case of equal variances.*

(c) *If lottery* L_y *is degenerate, then we obtain* $S(\Delta_{xy}) = S_x$.

Proof.   Straightforward calculations yield

$$\mathbb{E}\big[(\Delta_{xy} - \mathbb{E}[\Delta_{xy}])^3\big] = \mathbb{E}\big[([L_x - E_x] - [L_y - E_y])^3\big]$$
$$= \mathbb{E}\big[(L_x - E_x)^3\big] - 3\mathbb{E}\big[(L_x - E_x)^2(L_y - E_y)\big]$$
$$+ 3\mathbb{E}\big[(L_x - E_x)(L_y - E_y)^2\big] - \mathbb{E}\big[(L_y - E_y)^3\big],$$

which gives the above formula for $S(\Delta_{xy})$. Now Parts (a) and (c) directly follow, so we only need to prove Part (b). Suppose that L_x and L_y are stochastically independent. Then, we obtain

$$\mathbb{E}\big[(L_x - E_x)(L_y - E_y)^2\big] = \mathbb{E}\big[L_x L_y^2\big] - E_x \mathbb{E}\big[L_y^2\big]$$
$$= \mathbb{E}\big[L_x\big]\mathbb{E}\big[L_y^2\big] - E_x \mathbb{E}\big[L_y^2\big]$$
$$= 0,$$

where the first equality holds by $Cov(L_x, L_y) = 0$ and the second equality holds by $Cov(L_x, L_y^2) = 0$.   □

Next, we restrict attention to those binary lotteries that form a Mao pair (see Definition 7). The following lemma summarizes the properties of Mao pairs that we have already discussed in the main text and extends the list by some further useful

properties. In particular, the left-skewed lottery of a Mao pair is skewed relative to the right-skewed lottery only if the lotteries are positively correlated and not too skewed in absolute terms.

LEMMA A.2. *For any Mao pair M(E, V, S, η) the following statements hold:*

(a) *The covariance of L(E, V, −S) and L(E, V, S) strictly increases in η.*

(b) $Cov(L(E, V, −S), L(E, V, S)) > 0$ *if and only if* $\eta > (1 + S/\sqrt{4 + S^2})/2$.

(c) *The coskewness of L(E, V, −S) relative to L(E, V, S) strictly increases in η.*

(d) $Cos(L(E, V, S), L(E, V, −S)) = −Cos(L(E, V, −S), L(E, V, S)).$

(e) $Cos(L(E, V, −S), L(E, V, S)) > 0$ *if and only if* $\eta > (1 + S/\sqrt{4 + S^2})/2$.

(f) *The third standardized central moment of* $\Delta = L(E, V, −S) − L(E, V, S)$ *strictly increases in η.*

(g) *The lottery L(E, V, −S) is skewed relative to the lottery L(E, V, S) if and only if* $\eta > (2/3)(1 + S/\sqrt{4 + S^2})$. *In particular, for* $\eta = 1$, *L(E, V, −S) is skewed relative to L(E, V, S) if and only if* $S < (2/3)\sqrt{3} \approx 1.15$.

Proof. The proof is straightforward and therefore omitted.   □

A.2. *Independent Binary Lotteries*

Suppose the choice set is given by $\mathcal{C} = \{L(E, V, S_x), L(E, V, S_y)\}$ with $S_y > S_x$. In addition, let the lotteries be independent, in which case the relative skewness of the two lotteries is fully determined by the difference in their absolute skewness (see Lemma A.1 in Appendix A.1). In this sense, the choice between two independent binary lotteries with the same expected value and the same variance is structurally similar to the choice between a binary lottery and a safe option, as also in the latter case the relative skewness is given by the difference in absolute skewness, that is, simply by the skewness of the binary lottery (see Lemma A.1).

The first result in this section relates a salient thinker's choice between two independent binary lotteries with the same expected value and the same variance to her choice between a binary lottery and the safe option paying its expected value. Proposition A.1 shows, in particular, that a salient thinker does not necessarily choose the lottery that is more skewed in absolute terms if either both lotteries are right-skewed or both are left-skewed. This prediction is in line with experimental evidence by Ebert (2015).

PROPOSITION A.1. *Let* $\hat{S} \in \mathbb{R}$ *be the cutoff value derived in Proposition 3. Then, if the two binary lotteries are independent, the following statements hold:*

(a) $\lim\limits_{S_y \to \infty} \left[U^s\left(L(E, V, S_y) \,|\, \mathcal{C}\right) − U^s\left(L(E, V, S_x)|\mathcal{C}\right) \right] > 0 \iff S_x < \hat{S}.$

(b) $\lim\limits_{S_x \to −\infty} \left[U^s\left(L(E, V, S_y) \,|\, \mathcal{C}\right) − U^s\left(L(E, V, S_x)|\mathcal{C}\right) \right] > 0 \iff S_y > \hat{S}.$

Proof. Denote as π_{ij} the probability of state $s_{ij} = (x_i, y_j)$. First, we define

$$\lambda := \frac{\sqrt{p_x(1-p_x)p_y(1-p_y)}}{\sum_{s_{ij} \in S} \pi_{ij}\sigma(x_i, y_j)}.$$

In addition, let $p_z = p_z(S_z)$ be the marginal probability with which the lower payoff of the lottery $L_z := L(E, V, S_z)$ is realized. For the sake of brevity, we further write $U^s(L_z) = U^s(L_z \mid \mathcal{C})$. Then, by equation (2), we obtain

$$\frac{U^s(L_x) - E}{\lambda\sqrt{V}} = \sqrt{\frac{p_y}{1-p_y}}(\sigma(x_2, y_1) - \sigma(x_1, y_1))$$

$$+ \sqrt{\frac{1-p_y}{p_y}}(\sigma(x_2, y_2) - \sigma(x_1, y_2))$$

as well as

$$\frac{U^s(L_y) - E}{\lambda\sqrt{V}} = \sqrt{\frac{p_x}{1-p_x}}(\sigma(x_1, y_2) - \sigma(x_1, y_1))$$

$$+ \sqrt{\frac{1-p_x}{p_x}}(\sigma(x_2, y_2) - \sigma(x_2, y_1)).$$

Let $\mathrm{sgn} : \mathbb{R} \to \{-1, 0, 1\}$ be the signum-function. Taking the difference of the previous expressions yields

$$\mathrm{sgn}(U^s(L_y) - U^s(L_x)) = \mathrm{sgn}\bigg([\sigma(x_1, y_1) - \sigma(x_2, y_1)]\sqrt{p_x}\sqrt{p_y}$$

$$+ [\sigma(x_1, y_2) - \sigma(x_1, y_1)]p_x\sqrt{\frac{1-p_y}{1-p_x}}$$

$$+ [\sigma(x_1, y_2) - \sigma(x_2, y_2)](1-p_y)\sqrt{\frac{p_x}{p_y}}$$

$$+ [\sigma(x_2, y_2) - \sigma(x_2, y_1)]\sqrt{1-p_x}\sqrt{1-p_y}\bigg).$$

Now, as S_y approaches infinity, we obtain

$$\mathrm{sgn}\bigg(\lim_{S_y \to \infty}[U^s(L_y) - U^s(L_x)]\bigg)$$

$$= \mathrm{sgn}\bigg(\sqrt{p_x(1-p_x)}\sigma(x_1, E) - \sqrt{p_x(1-p_x)}\sigma(x_2, E)\bigg),$$

which exceeds zero if and only if $S_x < \hat{S}$. This yields part (a). Analogously,

$$\mathrm{sgn}\left(\lim_{S_x \to -\infty}\left[U^s(L_y) - U^s(L_x)\right]\right)$$
$$= \mathrm{sgn}\left(\sqrt{p_y(1-p_y)}\sigma(E, y_2) - \sqrt{p_y(1-p_y)}\sigma(E, y_1)\right),$$

which exceeds zero if and only if $S_y > \hat{S}$. This yields part (b).   □

Proposition A.1 suggests that under the assumption of independence a salient thinker prefers right- over left-skewed binary lotteries with the same expected value and the same variance. This prediction is consistent with evidence from the lab (Ebert and Wiesen 2011; Ebert 2015). More precisely, Proposition A.1(a) states that a salient thinker chooses an extremely right-skewed lottery $L(E, V, S_y)$ over the less skewed alternative $L(E, V, S_x)$ if and only if lottery $L(E, V, S_x)$ is *sufficiently* less skewed in the sense that $S_x < \hat{S}$, where the threshold value $\hat{S}$ is the same as the one derived in Proposition 3. This result follows from the fact that, by Corollary 2, a salient thinker's valuation of $L(E, V, S_y)$ approaches the lottery's expected value as S_y approaches infinity, so that in this limit case the choice between the two lotteries $L(E, V, S_y)$ and $L(E, V, S_x)$ is basically the same as the choice between the safe option E and the lottery $L(E, V, S_x)$. Analogously, Proposition A.1(b) implies that a salient thinker does *not* choose an extremely left-skewed lottery as long as the alternative option is sufficiently more skewed.

In summary, the preceding proposition implies that a lottery's absolute skewness matters, but it also suggests that more absolute skewness is not per se attractive to a salient thinker. By Proposition A.1(a) an extremely right-skewed option is *not* chosen over a less but still sufficiently skewed alternative, whereas by Proposition A.1(b) an extremely left-skewed lottery can be attractive relative to a *slightly* more skewed alternative. These predictions link back to Corollary 2, according to which the certainty equivalent to a binary lottery with a fixed expected value—when interpreted as a function of the lottery's skewness—is bounded. Notably, Ebert (2015) finds in his lab experiment that many subjects choose the less skewed of two right-skewed binary lotteries with the same expected value and variance. In this sense, Proposition A.1(a) is consistent with the experimental evidence by Ebert (2015) under the assumption that the subjects in his experiment perceived the lotteries as being independent.

Finally, by observing that a salient thinker's behavior is driven by the difference in absolute skewness, we can formulate the above result also in terms of relative skewness: Proposition A.1 basically says that there exists some threshold value $\hat{\Delta} \in \mathbb{R}$ such that the salient thinker chooses $L(E, V, S_y)$ over $L(E, V, S_x)$ if its relative skewness, $S(\Delta_{yx})$, exceeds the threshold $\hat{\Delta}$. It is in this sense that Proposition A.1 points towards a salient thinker's preference for relative skewness.

If we assume that one of the lotteries is symmetric, we further obtain the following corollary, which relates to the fourfold pattern of risk attitudes.

COROLLARY A.1.   *Suppose that the two lotteries are stochastically independent. Then, the following statements hold:*

  (a) *Let $S_x = 0$ and suppose that lottery $L(E, V, S_x)$ has non-negative payoffs. Then, there exists some $S' \in \mathbb{R}$ such that for any $S_y > S'$ the salient thinker chooses $L(E, V, S_y)$.*

  (b) *Let $S_y = 0$ and suppose that lottery $L(E, V, S_y)$ has non-positive payoffs. Then, there exists some $S'' \in \mathbb{R}$ such that for any $S_x < S''$ the salient thinker chooses $L(E, V, S_y)$.*

Recall from Corollary 1(a) that a binary lottery $L(E, V, S_x)$ with non-negative payoffs is chosen over its expected value if and only if $S_x > \hat{S} > 0$ holds. Then, it follows immediately from Proposition A.1(a) that a salient thinker chooses the more skewed alternative $L(E, V, S_y)$ if its absolute skewness exceeds a certain threshold (Part (a) of Corollary A.1). Using the same line of argumentation, Corollary A.1(b) follows directly from Corollary 1(b) and Proposition A.1(b). Again, both parts of Corollary A.1 suggest that a salient thinker opts for the lottery that is more skewed in absolute terms if it is also sufficiently skewed in relative terms.

A.3. A Preference for Relative Skewness can Explain the Allais Paradoxes

Not only choices on gambling, insurance, asset, and labor markets are crucially affected by skewness preferences, but also long-standing puzzles in choice under risk such as the *Allais paradoxes* can be attributed to skewness preferences or, more precisely, a preference for relative skewness.

Common-Consequence Allais Paradox.   In order to highlight the role that relative skewness plays in predicting whether subjects exhibit the common-consequence Allais paradox, we build on a recent experiment by Frydman and Mormann (2018). Suppose a subject chooses between the lotteries $L^1(z) = (25, 0.33; 0, 0.01; z, 0.66)$ and $L^2(z) = (24, 0.34; z, 0.66)$, where $z \in \{0, 24\}$. The common finding in the literature is that without making the state space explicit a majority of subjects choose $L^2(24)$ over $L^1(24)$, but $L^2(0)$ over $L^1(0)$. This preference reversal constitutes a puzzle from a classical economics point of view as according to EUT the common consequence z should not affect behavior. It can, however, be easily rationalized by CPT or salience theory. In addition, Frydman and Mormann (2018) have demonstrated that the emergence of the Allais paradox crucially depends on the lotteries' correlation structure. Can the idea of relative skewness explain this phenomenon?

In order to answer this question, let us start by considering the case in which the common consequence is given by $z = 24$. In this case, subjects choose between the safe option $L^2(24)$ and the left-skewed lottery $L^1(24) = (25, 0.33; 0, 0.01; 24, 0.66)$. According to the contrast effect, the left-skewed lottery's lowest payoff attracts a great deal of attention, which results in a preference for the safe option. Now suppose that the common consequence is given by $z = 0$. In this case, in order to derive predictions

TABLE A.1. Joint distribution of the common-consequence Allais paradox lotteries.

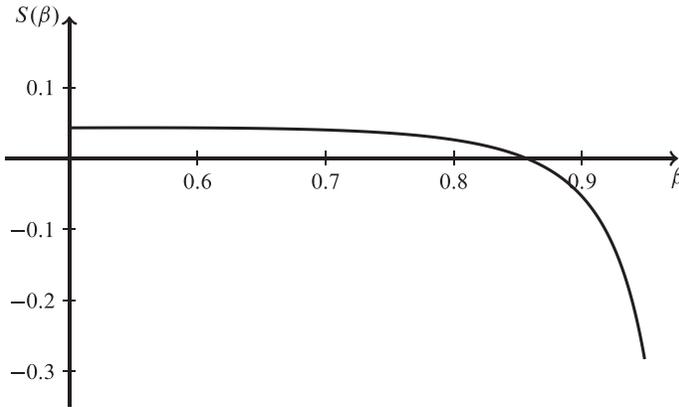| Probability	$0.33(2\beta - 1)$	$0.67 - 0.66\beta$	0.66β	$0.66(1 - \beta)$
$L^1(0)$	25	0	0	25
$L^2(0)$	24	24	0	0
$L^1(0) - L^2(0)$	1	−24	0	25

FIGURE A.1. The figure illustrates how skewed $L^1(0)$ is relative to $L^2(0)$ depending on β.

on a salient thinker's behavior, we need to state what the state space is. Adopting the notation by Frydman and Mormann (2018), we parameterize the lotteries' joint distribution via some $\beta \in [\frac{1}{2}, 1]$. The last row in Table A.1 depicts the difference in the outcomes of the lotteries, which determines their relative skewness.

Frydman and Mormann (2018) verify in their appendix that (i) the correlation between $L^1(0)$ and $L^2(0)$ strictly increases in $\beta \in [\frac{1}{2}, 1]$ and that (ii) a salient thinker's preference for $L^1(0)$ over $L^2(0)$ is diminished as the lotteries become more positively correlated. Moreover, as illustrated in Figure A.1, lottery $L^1(0)$ is the more skewed relative to lottery $L^2(0)$ the smaller β is, where we denote $S(\beta) := S(L^1(0) - L^2(0))$.[17] In this sense, salience theory again predicts a preference for relative skewness and it suggests that subjects are the more likely to exhibit the Allais paradox the more skewed lottery $L^1(0)$ is relative to lottery $L^2(0)$.

In their experiments, Frydman and Mormann (2018) vary the lotteries' correlation across three conditions—*no correlation* for $\beta = 0.67$, *intermediate correlation* for $\beta = 0.98$, and *maximal correlation* for $\beta = 1$—and they observe that, in line with salience, the share of subjects exhibiting the Allais paradox strictly decreases in β. Since we obtain $S(0.67) > S(0.98) > S(1)$, these findings can be understood as subjects revealing a preference for relative skewness.

---

17.   To be precise, $S(\beta)$ monotonically decreases in $\beta \in [0.56, 1]$, with $S(\beta) > 0$ if and only if $\beta < 0.86$.

TABLE A.2. Joint distribution of the common-ratio Allais paradox lotteries.

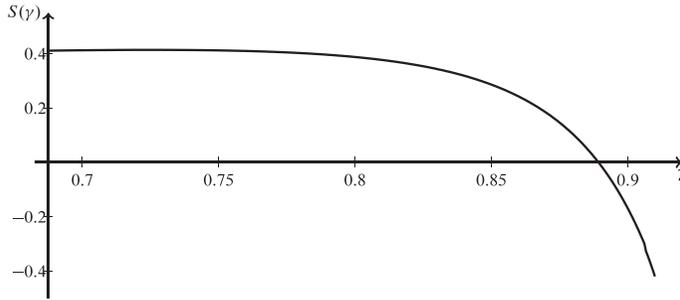| Probability	$0.8(1 - \gamma) - 0.05$	$0.25 - 0.8(1 - \gamma)$	$0.8(1 - \gamma)$	0.8γ
$L^3(0.25)$	40	40	0	0
$L^4(0.25)$	0	30	30	0
$L^3(0.25) - L^4(0.25)$	40	10	-30	0

FIGURE A.2. The figure illustrates how skewed $L^3(0.25)$ is relative to $L^4(0.25)$ depending on γ.

Common-Ratio Allais Paradox.   Finally, we demonstrate that also prominent versions of the common-ratio Allais paradox proposed by Kahneman and Tversky (1979) or O'Donoghue and Sprenger (2018) can be attributed to skewness preferences. We borrow the lotteries introduced in Problems 3 and 4 of Kahneman and Tversky (1979), but divide all payoffs by 1,000 for the sake of comparability with the common-consequence example. Let $L^3(q) = (40, 0.8q; 0, 1 - 0.8q)$ and $L^4(q) = (30, q; 0, 1 - q)$ for $q \in \{0.25, 1\}$. Kahneman and Tversky (1979) observe that around 80% of the subjects choose $L^4(1)$ over $L^3(1)$, which can be explained by an aversion toward left-skewed risks (see Proposition 3), whereas around 65% of the subjects choose $L^3(0.25)$ over $L^4(0.25)$. This preference reversal contradicts EUT, but is in line with Proposition A.1 given that the subjects perceived the lotteries as being independent. In order to say more on the role of relative skewness, Table A.2 parameterizes the joint distribution for $q = 0.25$ via some $\gamma \in [0.6875, 0.9375]$.

Note that for any salience function the states of the world can be unambiguously ranked according to salience: $\sigma(40, 0) > \sigma(30, 0) > \sigma(40, 30)$, where the first inequality follows by ordering and the second one by ordering and diminishing sensitivity. Then, it is easy to check that either the salient thinker always prefers $L^3(0.25)$ or she always prefers $L^4(0.25)$ or there exists some $\hat{\gamma} \in (0.6875, 0.9375]$ so that she prefers $L^3(0.25)$ over $L^4(0.25)$ if and only if $\gamma < \hat{\gamma}$. In addition, as illustrated in Figure A.2, lottery $L^3(0.25)$ is the more skewed relative to lottery $L^4(0.25)$ the smaller γ is,[18] where we denote $S(\gamma) := S(L^3(0.25) - L^4(0.25))$. As before, salience theory suggests that

---

18.   To be precise, $S(\gamma)$ monotonically decreases in $\gamma \in [0.7257, 0.9375]$, with $S(\gamma) > 0$ if and only if $\gamma < 0.89$.

subjects are the more likely to exhibit the Allais paradox the more skewed $L^3(0.25)$ is relative to $L^4(0.25)$. In this sense, also the common-ratio Allais paradox can be understood as a manifestation of skewness preferences.

Appendix B: Proofs

B.1. Auxiliary Results

In this section, we derive auxiliary results on the salience-weighted utility that we will apply in the proofs of our main results. In a first step, we characterize the ordering property via properties of the partial derivatives of the salience function (Lemma B.1). In a second step, we argue that the salience-weighted utility of any lottery with a finite expected utility is bounded (Lemma B.2). In a third step, we formally introduce the notion of first-order stochastic dominance and we state an implication for positive monotone transformations of lotteries that can be ordered in terms of first-order stochastic dominance (Lemma B.3). While Lemmata B.2 and B.3 are only relevant for the proof of Proposition 1, we will use Lemma B.1 also in the proof of Proposition 4.

The first result in this section pins down the sign of the partial derivatives of the salience function (whenever these partial derivatives exist). In particular, we show that, on any dense subset of $\mathbb{R}$, the partial derivatives of a salience function are different from zero.

LEMMA B.1. *Without loss of generality assume $x \geq y$. Consider a symmetric, continuous, and almost everywhere (a.e.) continuously differentiable function $\sigma :$ $\mathbb{R}^2 \to \mathbb{R}_+$. Denote as $N_x \subset \mathbb{R}$ the set on which the partial derivative $\partial\sigma(x, y)/\partial x$ does not exist and as $N_y \subset \mathbb{R}$ the set on which the partial derivative $\partial\sigma(x, y)/\partial y$ does not exist. Then, the following two statements are equivalent:*

(i) *The function σ satisfies ordering.*
(ii) *$\partial\sigma(x, y)/\partial x \geq 0$ for all $(x, y) \in \{\mathbb{R} \setminus N_x\} \times \mathbb{R}$ and $\partial\sigma(x, y)/\partial y \leq 0$ for all $(x, y) \in \mathbb{R} \times \{\mathbb{R} \setminus N_y\}$. In addition, we have $\partial\sigma(x, y)/\partial x > 0 > \partial\sigma(x, y)/\partial y$ on any dense subset of $\mathbb{R}$.*

Proof. (i)⟹(ii): For any $\varepsilon > 0$ ordering implies $\sigma(x + \varepsilon, y) - \sigma(x, y) > 0$. Thus, whenever the partial derivative $\partial\sigma(x, y)/\partial x$ exists, then $\partial\sigma(x, y)/\partial x \geq 0$ follows immediately from its definition. By the Mean Value Theorem, there exists some $\xi \in [x, x + \varepsilon]$ such that

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

Since N_x has measure zero by assumption, we obtain $\partial\sigma(x, y)/\partial x > 0$ on any dense subset of $\mathbb{R}$. The statement then follows by symmetry.

(ii)$\Rightarrow$(i): For any ε, $\varepsilon' \geq 0$ with $\varepsilon + \varepsilon' > 0$ it holds that

$$\sigma(x + \varepsilon, y - \varepsilon') - \sigma(x, y)$$

$$= \int_{[x,x+\varepsilon]\setminus N_x} \frac{\partial}{\partial z}\sigma(z, y - \varepsilon')dz - \int_{[y-\varepsilon',y]\setminus N_y} \frac{\partial}{\partial z}\sigma(x, z)dz > 0$$

since N_x and N_y have measure zero and either $[x, x + \varepsilon]\setminus N_x$ or $[y - \varepsilon', y]\setminus N_y$ or both are nonempty and include dense subsets of $\mathbb{R}$. Hence, σ satisfies ordering, which was to be proven. $\quad\square$

The second result in this section replaces the linear value function assumed in the main text by any strictly increasing value function $u(\,\cdot\,)$, and it states that the salience-weighted utility of any lottery with a finite expected utility (in particular, those with a finite expected value) is bounded.

LEMMA B.2.   *Fix a choice set* $\mathcal{C} := \{L_x, L_y\}$. *There exist constants* $\underline{K}, \bar{K} \in \mathbb{R}$ *so that* $U^s(L_z|\mathcal{C}) \in [\underline{K}, \bar{K}]$. *In particular, the bounds on the salience-weighted utility of lottery* L_z *are independent of lottery* L_{-z}.

Proof.   Since the expected utility of lottery $L_z \in \mathcal{C}$ is finite, the statement follows from the fact that the salience function is bounded (away from zero). $\quad\square$

Before we state the last result of this section, we formally introduce the notion of first-order stochastic dominance: we say that *a lottery* L_x *first-order stochastically dominates the lottery* L_y *(i.e.,* L_x *fosd* L_y*) if and only if* $F_x(z) \leq F_y(z)$ *for any* $z \in \mathbb{R}$ *and* $F_x(z) < F_y(z)$ *for some* $z \in \mathbb{R}$. Lemma B.3 states that not only has a first-order stochastic dominant lottery a higher expected value than the corresponding dominated lottery, but that this inequality survives under any positive monotone transformation of the lotteries.

LEMMA B.3.   *Let* $\varphi : \mathbb{R} \to \mathbb{R}$ *be a strictly increasing function with* $\varphi'(\cdot) > 0$ *almost everywhere. Then, it holds*

$$L_x \text{ fosd } L_y \quad \Longrightarrow \quad \int_{\mathbb{R}} \varphi(x)\, dF_x > \int_{\mathbb{R}} \varphi(y)\, dF_y.$$

Proof.   The proof is straightforward and therefore omitted. $\quad\square$

B.2. Main Results

Proof of Proposition 1.   We prove a slightly more general version of this proposition. Indeed, the statement holds not only for a linear value function, but for any strictly increasing value function $u(\cdot)$ with $u'(\cdot) > 0$ such that the expected utility, $U(L) = \int_{\mathbb{R}} u(x)\, dF(x)$, is finite.

To begin with, notice that a salient thinker with a value function $u(\cdot)$ strictly prefers the lottery L to the safe option c if and only if

$$u(c) < \int_{\mathbb{R}} u(x) \cdot \frac{\sigma(u(x), u(c))}{\int_{\mathbb{R}} \sigma(u(s), u(c)) \, dF(s)} \, dF(x) =: h(c).$$

The proof now proceeds in four steps. In a first step, we use the Intermediate Value Theorem and Lemma B.2 to argue that $u(c) = h(c)$ has at least one solution, which implies that a certainty equivalent exists. In a second step, we show that for any c that solves $u(c) = h(c)$, there exists some $\varepsilon > 0$ such that $u(c') > h(c')$ for any $c' \in (c, c + \varepsilon)$ and $u(c'') < h(c'')$ for any $c'' \in (c - \varepsilon, c)$. This implies that the certainty equivalent is *locally* unique. In a third step, we take two lotteries L_x and L_y where L_x *fosd* L_y and show—by the use of Lemma B.3—that the smallest certainty equivalent to the lottery L_x strictly exceeds the largest certainty equivalent to the lottery L_y. In a fourth step, we combine the results from the previous steps to prove that the certainty equivalent is (globally) unique. Monotonicity then follows immediately from the third step.

Step 1.   Since both $u(\cdot)$ and $h(\cdot)$ are continuous, the Intermediate Value Theorem states that at least one solution to $u(c) = h(c)$ exists if it is true that

$$\lim_{c \to -\infty} u(c) < \lim_{c \to -\infty} h(c) \quad \text{and} \quad \lim_{c \to \infty} u(c) > \lim_{c \to \infty} h(c). \tag{B.1}$$

In order to verify that equation (B.1) is indeed fulfilled, we distinguish three cases.

First, suppose that $\lim_{c \to -\infty} u(c) = -\infty$ and $\lim_{c \to \infty} u(c) = \infty$. By Lemma B.2, there exist constants $\underline{K}, \overline{K} \in \mathbb{R}$ such that $h(c) \in [\underline{K}, \overline{K}]$ for any $c \in \mathbb{R}$. Hence, equation (B.1) holds.

Second, let both $\lim_{c \to -\infty} u(c)$ and $\lim_{c \to \infty} u(c)$ be finite. Since $u'(\cdot) > 0$,

$$\lim_{z \to -\infty} u(z) = \int_{\mathbb{R}} \frac{\sigma(u(x), u(c))}{\int_{\mathbb{R}} \sigma(u(s), u(c)) \, dF(s)} \cdot \lim_{z \to -\infty} u(z) \, dF(x)$$

$$< \int_{\mathbb{R}} \frac{\sigma(u(x), u(c))}{\int_{\mathbb{R}} \sigma(u(s), u(c)) \, dF(s)} \cdot u(x) \, dF(x)$$

$$= h(c),$$

for any $c \in \mathbb{R}$. This implies, in particular, that $\lim_{c \to -\infty} u(c) < \lim_{c \to -\infty} h(c)$ has to hold. By the same type of argument, we obtain $\lim_{z \to \infty} u(z) > h(c)$ for any $c \in \mathbb{R}$, which, in turn, implies that $\lim_{c \to \infty} u(c) > \lim_{c \to \infty} h(c)$ has to hold. Hence, equation (B.1) is satisfied.

Third, suppose that *either* $\lim_{c \to -\infty} u(c) = -\infty$ and $\lim_{c \to \infty} u(c) \in \mathbb{R}$ *or* $\lim_{c \to -\infty} u(c) \in \mathbb{R}$ and $\lim_{c \to \infty} u(c) = \infty$. Here, combining the arguments used in the first and second case yields the claim. Altogether, we conclude that at least one certainty equivalent exists.

Step 2.   Next, consider some certainty equivalent c. We show that there exists some $\varepsilon > 0$ such that $u(c') > h(c')$ for any $c' \in (c, c + \varepsilon)$ and $u(c'') < h(c'')$ for any

$c'' \in (c - \varepsilon, c)$. Since we have $u'(c) > 0$ by assumption, it is sufficient to verify that $h'(c) \leq 0$ holds.

Denote as $N_c \subset \mathbb{R}$ the set on which the partial derivative $\partial \sigma(u(x), u(c))/\partial c$ does not exist. Since N_c has measure zero, we know that $h'(c)$ exists. Moreover, as σ is bounded and as $|u|$ is integrable, the Dominated Convergence Theorem implies that we can reverse the order of differentiation and integration. Simple re-arrangements show that $h'(c) \leq 0$ holds if and only if

$$\underbrace{\frac{\int_{\mathbb{R}} u(x)\sigma(u(x), u(c)) \, dF(x)}{\int_{\mathbb{R}} \sigma(u(s), u(c)) \, dF(s)}}_{=h(c)} \left(\int_{\mathbb{R} \setminus N_c} \frac{\partial}{\partial c}\sigma(u(s), u(c)) \, dF(s) \right)$$

$$\geq \int_{\mathbb{R} \setminus N_c} u(x) \frac{\partial}{\partial c}\sigma(u(x), u(c)) \, dF(x).$$

Let $\underline{X} := \{x \in \mathbb{R} : u(x) \leq u(c)\} \setminus N_c$ and $\overline{X} := \{x \in \mathbb{R} : u(x) > u(c)\} \setminus N_c$. Since c is a certainty equivalent by assumption, we have $h(c) = u(c)$ and the preceding inequality is equivalent to

$$\int_{\underline{X}} \underbrace{(u(c) - u(x))}_{\geq 0} \underbrace{\frac{\partial}{\partial c}\sigma(u(x), u(c))}_{\geq 0} \, dF(x)$$

$$+ \int_{\overline{X}} \underbrace{(u(c) - u(x))}_{<0} \underbrace{\frac{\partial}{\partial c}\sigma(u(x), u(c))}_{\leq 0} \, dF(x) \geq 0,$$

where the sign of $\partial \sigma(u(x), u(c))/\partial c = u'(c) \cdot \partial \sigma(u(x), u(c))/\partial u(c)$ follows from Lemma B.1. As a consequence, we have $h'(c) \leq 0$, which implies that any certainty equivalent c is locally unique.

Step 3.   Consider two lotteries L_x and L_y where L_x *fosd* L_y. Suppose that $c_z \in \mathbb{R}$ denotes some certainty equivalent to lottery $L_z \in \{L_x, L_y\}$. Then any such certainty equivalent c_z solves

$$\int_{\mathbb{R}} \varphi(s, c_z) \, dF_z(s) = 0, \tag{B.2}$$

where $\varphi(s, c) := (u(s) - u(c))\sigma(u(s), u(c))$ and F_z is the cumulative distribution function of L_z.

Since ordering (by Lemma B.1) implies that

$$\frac{\partial}{\partial s}\sigma(u(s), u(c)) \geq 0 \text{ (a.e.)} \quad \text{if and only if} \quad u(s) \geq u(c),$$

$\partial \varphi(s, c)/\partial s > 0$ almost everywhere. Analogously, $\partial \varphi(s, c)/\partial c < 0$ almost everywhere.

Now, for the sake of a contradiction, let $c_x \leq c_y$. This assumption implies

$$0 = \int_{\mathbb{R}} \varphi(s, c_x)\, dF_x(s) \geq \int_{\mathbb{R}} \varphi(s, c_y)\, dF_x(s) > \int_{\mathbb{R}} \varphi(s, c_y)\, dF_y(s) = 0,$$

where the equalities follow from (B.2), the weak inequality follows from $\partial \varphi(s, c)/\partial c$ < 0 (a.e.) and $c_x \leq c_y$, and the strict inequality holds by $\partial \varphi(s, c)/\partial s > 0$ (a.e.) and Lemma B.3; a contradiction.

Hence, we conclude that the smallest certainty equivalent to the lottery L_x is larger than the largest certainty equivalent to the lottery L_y.

Step 4. Consider some lottery L with a cumulative distribution function F. For the sake of a contradiction, suppose that at least two certainty equivalents to L exist; that is, there exist some $c_1, c_2 \in \mathbb{R}$, $c_1 < c_2$, such that $u(c_1) = h(c_1)$ and $u(c_2) = h(c_2)$.

Now consider a sequence of lotteries $(L_n)_{n \in \mathbb{N}}$ with cumulative distribution functions $(F_n)_{n \in \mathbb{N}}$ such that L *fosd* L_n for any $n \in \mathbb{N}$ and F_n converges to F pointwise. By the second step, for each $k \in \{1, 2\}$ there is some $\varepsilon_k > 0$ such that $u(z) > h(z)$ for any $z \in (c_k, c_k + \varepsilon_k)$ and $u(z) < h(z)$ for any $z \in (c_k - \varepsilon_k, c_k)$. In addition, it is straightforward to see that

$$\int_{\mathbb{R}} \frac{u(x)\sigma(u(x), u(c))}{\int_{\mathbb{R}} \sigma(u(s), u(c))\, dF_n(s)}\, dF_n(x) \overset{n \to \infty}{\longrightarrow} \int_{\mathbb{R}} \frac{u(x)\sigma(u(x), u(c))}{\int_{\mathbb{R}} \sigma(u(s), u(c))\, dF(s)}\, dF(x).$$

Together, these two observations imply that there exists some $n' \in \mathbb{N}$ such that for any $n \geq n'$ also L_n has at least two certainty equivalents c_1^n and c_2^n that smoothly converge to c_1 and c_2, respectively, when n approaches infinity. This implies that there exists some $n'' \geq n'$ such that for any $n \geq n''$ we have $c_2^n > c_1$. But this yields a contradiction to the fact that the smallest certainty equivalent to L is larger than the largest certainty equivalent to L_n (see Step 3). Hence, the certainty equivalent is also globally unique. $\qed$

Proof of Proposition 2. Let $\mathcal{C} = \{L_x, L_y\}$ and suppose that L_x and L_y are stochastically independent. Again, we prove a slightly more general version, allowing for any increasing value function $u(\cdot)$ with $u'(\cdot) > 0$ and finite expected utility, $U(L_z) = \int_{\mathbb{R}} u(z)\, dF(z, w) < \infty$. Denote the marginal cumulative distribution functions as F_x and F_y, respectively. And, as before, we define $\varphi(x, y) := (u(x) - u(y))\sigma(u(x), u(y))$. Now suppose that L_x first-order stochastically dominates L_y and denote $\Phi(x, L_y) := \int_{\mathbb{R}} \varphi(x, y)\, dF_y(y)$. As, for any $y \in \mathbb{R}$, $\partial\varphi(x, y)/\partial x > 0$ (a.e.), we also have $\partial\Phi(x, L_y)/\partial x > 0$. In addition, it is easy to check that, for any lottery L_y, we obtain $\int_{\mathbb{R}} \Phi(s, L_y)\, dF_y(s) = 0$. We thus conclude that

$$\int_{\mathbb{R}^2} \varphi(x, y)\, dF(x, y) = \int_{\mathbb{R}} \Phi(s, L_y)\, dF_x(s) > \int_{\mathbb{R}} \Phi(s, L_y)\, dF_y(s) = 0,$$

where the first equality holds by independence, and the inequality holds by $\partial\Phi(x, L_y)/\partial x > 0$ and Lemma B.3. By definition, $U^s(L_x \mid \mathcal{C}) > U^s(L_y \mid \mathcal{C})$ if and only if $\int_{\mathbb{R}^2} \varphi(x, y)\, dF(x, y) > 0$, which yields the claim. $\qed$

Proof of Proposition 3. Consider a binary lottery L with expected value E and variance V. Using the characterization of binary risks in equation (2), we observe that the lottery's lower payoff becomes more likely if the lottery's skewness increases. Formally, we have

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

Using (2) again, we conclude that—for a fixed expected value and a fixed variance—both the lower payoff $x_1 = x_1(E, V, S)$ and the higher payoff $x_2 = x_2(E, V, S)$ increase in the lottery's skewness S. Therefore, the difference between the lower (higher) payoff and the expected value monotonically decreases (increases) in the lottery's skewness S:

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

Since the expected value E is fixed, an increase in the contrast $|x_k - E|$ is equivalent to an increase in the salience of state (x_k, E) due to the ordering property. Hence, the lower payoff's salience decreases in S, whereas the higher payoff's salience increases in S.

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

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

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

Proof of Proposition 4. Let either $x_1 = x_1(E, V, S) \geq 0$ or $x_2 = x_2(E, V, S) \leq 0$. By definition, the threshold value $\hat{S} = \hat{S}(E, V)$ solves $r = r(E, V, S) = 0$. We proceed in two steps. First, we determine the sign of $\partial \hat{S}(E, V)/\partial E$ given that this partial derivative exists. Second, we argue that $\partial \hat{S}(E, V)/\partial E$ exists for any $E \in \mathbb{R} \setminus N$ where N has measure zero. This suffices to prove our claim.

Step 1. If $\partial \hat{S}(E, V)/\partial E$ exists, then the Implicit Function Theorem yields

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

For any expected value E, variance V, and skewness S, we have

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

so that the sign of the preceding derivative is equal to the sign of

$$\frac{d}{dE}\left(\frac{\sigma(x_1, E) - \sigma(x_2, E)}{p\sigma(x_1, E) + (1-p)\sigma(x_2, E)}\right)$$

$$= \frac{\sigma(x_2, E)\frac{d}{dE}\sigma(x_1, E) - \sigma(x_1, E)\frac{d}{dE}\sigma(x_2, E)}{\left(p\sigma(x_1, E) + (1-p)\sigma(x_2, E)\right)^2}.$$

Thus, by definition of the decreasing level effect, we obtain $\partial r(E, V, S)/\partial E < 0$.

In addition, for any expected value E, and any variance V, we have

$$\left.\frac{\partial}{\partial S}r(E, V, S)\right|_{S=\widehat{S}} = \sqrt{V}\underbrace{\left(\frac{\hat{\sigma}_1 - \hat{\sigma}_2}{p\hat{\sigma}_1 + (1-p)\hat{\sigma}_2}\right)}_{=0 \text{ by definition of } \widehat{S}} \cdot \left.\frac{\partial}{\partial S}\left(\sqrt{p(1-p)}\right)\right|_{S=\widehat{S}}$$

$$+ \sqrt{Vp(1-p)} \cdot \left.\frac{\partial}{\partial S}\left(\frac{\sigma_1 - \sigma_2}{p\sigma_1 + (1-p)\sigma_2}\right)\right|_{S=\widehat{S}},$$

where $\sigma_k := \sigma(x_k(E, V, S), E)$ and $\hat{\sigma}_k := \sigma(x_k(E, V, \widehat{S}), E)$ for $k \in \{1, 2\}$. Since $\hat{\sigma}_1 = \hat{\sigma}_2$ by the definition of $\widehat{S}$ and since $\partial\widehat{S}(E, V)/\partial E$ exists by assumption, it is easy to verify that $\partial r(E, V, S)/\partial S|_{S=\widehat{S}}$ has the same sign as

$$\left.\frac{\partial}{\partial S}\left(\sigma(x_1(E, V, S), E) - \sigma(x_2(E, V, S), E)\right)\right|_{S=\widehat{S}}.$$

As $\partial\widehat{S}(E, V)/\partial E$ exists by assumption, it follows, by Lemma B.1 and equation (2),

$$\left.\frac{\partial}{\partial S}\left(\sigma(x_1(E, V, S), E) - \sigma(x_2(E, V, S), E)\right)\right|_{S=\widehat{S}} < 0.$$

This implies that $\partial\widehat{S}(E, V)/\partial E < 0$, whenever this partial derivative exists.

Step 2. Notice that

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

holds if and only if $\partial\sigma(x_k, E)/\partial x_k = 0$ for $k \in \{1, 2\}$. For a fixed variance V and a fixed skewness S, we define

$$N := \left\{E \in \mathbb{R} : \frac{\partial}{\partial x_k}\sigma(x_k(E, V, S), E) = 0, \; k \in \{1, 2\},\right.$$

$$\left. \text{or } \frac{d}{dE}\sigma(x_k(E, V, S), E) \text{ does not exist}\right\}.$$

Hence for any expected value $E \in \mathbb{R} \setminus N$ the partial derivative $\partial\widehat{S}(E, V)/\partial E$ exists. By Lemma B.1, the set N has measure zero, which implies that for any $\varepsilon > 0$ it holds that

$$\widehat{S}(E + \varepsilon, V) - \widehat{S}(E, V) = \int_{[E, E+\varepsilon]\setminus N} \frac{\partial}{\partial x}\widehat{S}(x, V)\,dx < 0.$$

This completes the proof.   □

Proof of Proposition 5. Let $L_y := L(E, V, S)$ and $L_x := L(E, V, -S)$, and denote as $p = p(-S)$ the probability with which the left-skewed lottery's lower payoff is realized. In addition, denote as sgn $: \mathbb{R} \to \{-1, 0, 1\}$ the signum-function. For the sake of brevity, we again write $U^s(L_z) = U^s(L_z|\mathcal{C})$.

We prove Parts (a), (b), and (c) successively. $\square$

Part (a). Since

$$\text{sgn}\big(U^s(L_y) - U^s(L_x)\big)$$

$$= \text{sgn}\bigg((1 - p)\Big[2\sigma(y_2, x_1) + \eta\big(\sigma(y_1, x_1) - 2\sigma(y_2, x_1) + \sigma(y_2, x_2)\big)\Big]$$

$$- \Big[2(1 - p)\sigma(y_1, x_2) + \eta\big(\sigma(y_1, x_1) - 2\sigma(y_2, x_1) + \sigma(y_2, x_2)\big)\Big]\bigg),$$

we conclude that $U^s(L_y) - U^s(L_x) > 0$ holds if and only if

$$\eta < \frac{2(1 - p)[\sigma(y_2, x_1) - \sigma(y_1, x_2)]}{2[(1 - p)\sigma(y_2, x_1) - p\sigma(y_1, x_2)] - (1 - 2p)[\sigma(y_1, x_1) + \sigma(y_2, x_2)]} =: \tilde{\eta}(S),$$

where $x_k = x_k(E, V, S)$, $y_k = y_k(E, V, S)$, $k \in \{1, 2\}$, and $p = p(-S)$ are defined in equation (2). Now we define $\check{\eta}(S) := \min\{1, \tilde{\eta}(S)\}$. Then, it is straightforward to check that $\check{\eta}(S) < 1$ if and only if

$$\sigma(y_1, x_1) + \sigma(y_2, x_2) - 2\sigma(y_1, x_2) < 0. \tag{B.3}$$

As $\lim_{S \to \infty} x_1 = -\infty$, as $\lim_{S \to \infty} x_2 = E = \lim_{S \to \infty} y_1$, and as $\lim_{S \to \infty} y_2 = \infty$, the ordering property implies that Inequality (B.3) does *not* hold in the limit of $S \to \infty$. Analogously, as $\lim_{S \to 0} x_1 = E - \sqrt{V} = \lim_{S \to 0} y_1$ and as $\lim_{S \to 0} x_2 = E + \sqrt{V} = \lim_{S \to 0} y_2$, ordering implies that Inequality (B.3) holds in the limit of $S \to 0$.

Part (b). Since we have

$$\text{sgn}\big(U^s(L_y) - U^s(L_x)\big) = \text{sgn}\big(p(y_2 - x_1)\sigma(y_2, x_1) + (1 - p)(y_1 - x_2)\sigma(y_1, x_2)\big)$$

$$= \text{sgn}\big(p(y_2 - x_1)\big(\sigma(y_2, x_1) - \sigma(y_1, x_2)\big)\big),$$

where the second equality follows from the fact that $\mathbb{E}[L_x] = E = \mathbb{E}[L_y]$, and since $y_2 > x_2 > y_1 > x_1$ by equation (2), the statement follows from ordering.

Part (c). Notice that y_1 and y_2 monotonically increase in S, whereas x_1 and x_2 monotonically decrease in S. Thus, since $y_2 > x_2 > y_1 > x_1$, ordering implies that $\sigma(y_1, x_1)$ and $\sigma(y_2, x_2)$ monotonically increase in S whereas $\sigma(y_1, x_2)$ monotonically decreases in S. Together these observations imply that the left-hand side of (B.3) strictly increases in S. The statement then follows from $\lim_{S \to \infty} \check{\eta}(S) = 1$ and $\lim_{S \to 0} \check{\eta}(S) < 1$ (see Part (a) of the proof).

Appendix C: Decreasing Level Effect

In this section, we show that a wide class of salience functions satisfies the decreasing level effect, as introduced in Definition 2. Consider, for instance, the following class of salience functions:

$$\sigma_\theta(x, y) := \frac{(x - y)^{2n}}{(|x| + |y| + \theta)^{2n}}, \, n \in \mathbb{N}, \, \theta > 0.$$

For the sake of the argument, let $y, z \in \mathbb{R}$ and $x \geq \max\{-y, -z, 0\}$. Then,

$$\frac{\frac{d}{dx}\sigma_\theta(x + y, x + z)^n}{\sigma_\theta(x + y, x + z)^n} = -\frac{4n}{x + y + x + z + \theta} = \frac{\frac{d}{dx}\sigma_\theta(-x - y, -x - z)^n}{\sigma_\theta(-x - y, -x - z)^n},$$

which implies that, for any $n \in \mathbb{N}$, the salience function $\sigma_\theta(x, y)$ indeed satisfies the decreasing level effect. It is easy to check that also positive transformations of the salience function $\sigma_\theta(x, y)$ satisfy this property.

Similar calculations show that also the salience function

$$\sigma(x, y) = \frac{|x - y|}{|x| + |y| + \theta},$$

which was proposed by Bordalo et al. (2012), satisfies the decreasing level effect, and so do the weighting functions proposed by Bordalo et al. (2016) or Thakral and To (2018).

References

Azevedo, Eduardo M. and Daniel Gottlieb (2012). "Risk-neutral Firms can Extract Unbounded Profits from Consumers with Prospect Theory Preferences." *Journal of Economic Theory*, 147, 1291–1299.

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

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw (2011). "Maxing Out: Stocks as Lotteries and the Cross-section of Expected Returns." *Journal of Financial Economics*, 99, 427–446.

Barberis, Nicholas C. (2013). "Thirty Years of Prospect Theory in Economics: A Review and Assessment." *Journal of Economic Perspectives*, 27(1), 173–195.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C. Teitelbaum (2013). "The Nature of Risk Preferences: Evidence from Insurance Choices." *American Economic Review*, 103 (6), 2499–2529.

Bell, David E. (1982). "Regret in Decision Making under Uncertainty." *Operations Research*, 30, 961–981.

Berkhout, Peter, Joop Hartog, and Dinand Webbink (2010). "Compensation for Earnings Risk under Worker Heterogeneity." *Southern Economic Journal*, 76, 762–790.

Bernoulli, Daniel (1738). "Specimen Theoriae Novae de Mensura Sortis." *Comentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, 175–192 [translated by L. Sommer (1954) in *Econometrica*, 22, 23–36].

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

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013a). "Salience and Asset Prices." *American Economic Review, Papers & Proceedings*, 103, 623–628.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013b). "Salience and Consumer Choice." *Journal of Political Economy*, 121, 803–843.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2016). "Competition for Attention." *Review of Economic Studies*, 83, 481–513.

Boyer, Brian, Todd Mitton, and Keith Vorkink (2010). "Expected Idiosyncratic Skewness." *Review of Financial Studies*, 23, 169–202.

Brocas, Isabelle, Juan Carrillo, Aleksandar Giga, and Fernando Zapatero (2016). "Skewness Seeking in a Dynamic Portfolio Choice Experiment." Available at https://www.cemfi.es/ftp/pdf/papers/wshop/skweness_march17.pdf.

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

Bruhin, Adrian, Maha Manai, and Luis Santos-Pinto (2019). "Risk and Rationality: The Relative Importance of Probability Weighting and Choice Set Dependence." Available at http://www.hec.unil.ch/lspinto/Papers%20&%20CV/PrWeighting_vs_ChoiceSetDependence.pdf.

Brunnermeier, Markus K. and Jonathan A. Parker (2005). "Optimal Expectations." *American Economic Review*, 95(4), 1092–1118.

Bushong, Ben, Matthew Rabin, and Joshua Schwartzstein (2017). "A Model of Relative Thinking." Available at https://www.hbs.edu/faculty/Publication%20Files/RelativeThinking_6a962ed2-9254-43d6-984d-6781ffe8a78e.pdf.

Campbell, John Y. and Robert J. Shiller (1988). "The Dividend-price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1, 195–228.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein (2001). "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices." *Journal of Financial Economics*, 61, 345–381.

Choi, Darwin, Dong Lou, and Abhiroop Mukherjee (2016). "The Effect of Superstar Firms on College Major Choice." Available at http://www.pbcsf.tsinghua.edu.cn/research/loudong/paper/W%20The%20Effect%20of%20Superstar%20Firms%20on.pdf.

Chunhachinda, Pornchai, Krishnan Dandapani, Shahid Hamid, and Arun J. Prakash (1997). "Portfolio Selection and Skewness: Evidence from International Stock Markets." *Journal of Banking & Finance*, 21, 143–167.

Conrad, Jennifer, Robert F. Dittmar, and Eric Ghysels (2013). "Ex Ante Skewness and Expected Stock Returns." *Journal of Finance*, 68, 85–124.

Dertwinkel-Kalt, M., M. Lange, K. Köhler, and T. Wenzel (2017). "Demand Shifts due to Salience Effects: Experimental Evidence." *Journal of the European Economic Association*, 15, 626–653.

Dertwinkel-Kalt, Markus and Mats Köster (2017). "Salient Compromises in the Newsvendor Game." *Journal of Economic Behavior & Organization*, 141, 301–315.

Dertwinkel-Kalt, Markus and Mats Köster (2019). "Optimal Stopping in a Dynamic Salience Model." Working paper.

Diecidue, Enrico and Jeeva Somasundaram (2017). "Regret Theory: A New Foundation." *Journal of Economic Theory*, 172, 88–119.

Ebert, Sebastian (2013). "Moment Characterization of Higher-order Risk Preferences." *Theory and Decision*, 74, 267–284.

Ebert, Sebastian (2015). "On Skewed Risks in Economic Models and Experiments." *Journal of Economic Behavior & Organization*, 112, 85–97.

Ebert, Sebastian and Christian Hilpert (2016). "Skewness Preference and the Popularity of Technical Analysis." Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2354962.

Ebert, Sebastian and Philipp Strack (2015). "Until the Bitter End: On Prospect Theory in a Dynamic Context." *American Economic Review*, 105(4), 1618–1633.

Ebert, Sebastian and Philipp Strack (2018). "Never, Ever Getting Started: On Prospect Theory Without Commitment." Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765550.

Ebert, Sebastian and Daniel Wiesen (2011). "Testing for Prudence and Skewness Seeking." *Management Science*, 57, 1334–1349.

Eeckhoudt, Louis and Harris Schlesinger (2006). "Putting Risk in its Proper Place." *American Economic Review*, 96(1), 280–289.

Enke, Benjamin and Florian Zimmermann (2019). "Correlation Neglect in Belief Formation." *Review of Economic Studies*, 86, 313–332.

Fama, Eugene F. and Kenneth R. French (1992). "The Cross-section of Expected Stock Returns." *Journal of Finance*, 47, 427–465.

Fischbacher, Urs (2007). "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10, 171–178.

Forrest, David, Robert Simmons, and Neil Chesters (2002). "Buying a Dream: Alternative Models of Demand for Lotto." *Economic Inquiry*, 40, 485–496.

Frydman, Cary and Milica Milosavljevic Mormann (2018). "The Role of Salience in Choice Under Risk: An Experimental Investigation." Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2778822.

Garrett, Thomas A. and Russell S. Sobel (1999). "Gamblers Favor Skewness, not Risk: Further Evidence from United States' Lottery Games." *Economics Letters*, 63, 85–90.

Genest, Wilfried, William R. Stauffer, and Wolfram Schultz (2016). "Utility Functions Predict Variance and Skewness Risk Preferences in Monkeys." *Proceedings of the National Academy of Sciences*, 113, 8402–8407.

Golec, Joseph and Maurry Tamarkin (1998). "Bettors Love Skewness, Not Risk, at the Horse Track." *Journal of Political Economy*, 106, 205–225.

Gollier, Christian (2018). "Aversion to Risk of Regret and Preference for Positively Skewed Risks." *Economic Theory*.

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

Green, T. Clifton and Byoung-Hyoun Hwang (2012). "Initial Public Offerings as Lotteries: Skewness Preference and First-day Returns." *Management Science*, 58, 432–444.

Greiner, Ben (2015). "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." *Journal of the Economic Science Association*, 1, 114–125.

Grossman, Philip J. and Catherine C. Eckel (2015). "Loving the Long Shot: Risk Taking with Skewed Lotteries." *Journal of Risk and Uncertainty*, 51, 195–217.

Grove, Wayne A., Michael Jetter, and Kerry L. Papps (2018). "Career Lotto: Labor Supply in Winner-Take-All Markets." IZA DP No. 12012.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2018). "Time Varying Risk Aversion." *Journal of Financial Economics*, 128, 403–421.

Hartog, Joop and Wim P. M. Vijverberg (2007). "On Compensation for Risk Aversion and Skewness Affection in Wages." *Labour Economics*, 14, 938–956.

Huber, Joel, John W. Payne, and Christopher Puto (1982). "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis." *Journal of Consumer Research*, 9, 90–98.

Kahneman, D. and A. Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47, 263–291.

Kőszegi, Botond and Adam Szeidl (2013). "A Model of Focusing in Economic Choice." *Quarterly Journal of Economics*, 128, 53–104.

Kontek, Krzysztof (2016). "A Critical Note on Salience Theory of Choice under Risk." *Economics Letters*, 149, 168–171.

Lanzani, Giacomo (2018). "An Axiomatic Approach to Salience Theory." Working paper.

Levy, Gilat and Ronny Razin (2015). "Correlation Neglect, Voting Behavior, and Information Aggregation." *American Economic Review*, 105(4), 1634–1645.

Loomes, Graham and Robert Sugden (1982). "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *The Economic Journal*, 92, 805–824.

Loomes, Graham and Robert Sugden (1987). "Some Implications of a More General Form of Regret Theory." *Journal of Economic Theory*, 41, 270–287.

Mao, James C. T. (1970). "Survey of Capital Budgeting: Theory and Practice." *Journal of Finance*, 25, 349–360.

Menezes, C. F., C. Geiss, and J. Tressler (1980). "Increasing Downside Risk." *American Economic Review*, 70(5), 921–932.

Mitton, Todd and Keith Vorkink (2007). "Equilibrium Underdiversification and the Preference for Skewness." *Review of Financial Studies*, 20, 1255–1288.

O'Donoghue, Ted and Charles Sprenger (2018). "Reference-Dependent Preferences." *Handbook of Behavioral Economics: Applications and Foundations 1*, Chapter 1, 1–77.

Prakash, Arun J., Chun-Hao Chang, and Therese E. Pactwa (2003). "Selecting a Portfolio with Skewness: Recent Evidence from US, European, and Latin American Equity Markets." *Journal of Banking & Finance*, 27, 1375–1390.

Quiggin, J. (1982). "A Theory of Anticipated Utility." *Journal of Economic Behavior & Organization*, 3, 323–343.

Rieger, Marc Oliver and Mei Wang (2006). "Cumulative Prospect Theory and the St. Petersburg Paradox." *Economic Theory*, 28, 665–679.

Schkade, D. A. and D. Kahneman (1998). "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction." *Psychological Science*, 9, 340–346.

Soltani, Alireza, Benedetto De Martino, and Colin Camerer (2012). "A Range-normalization Model of Context-dependent Choice: A New Model and Evidence." *PLoS Computational Biology*, 8, e1002607.

Strack, Philipp and Paul Viefers (2017). "Too Proud to Stop: Regret in Dynamic Decisions." Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2465840.

Strait, Caleb E. and Benjamin Y. Hayden (2013). "Preference Patterns for Skewed Gambles in Rhesus Monkeys." *Biology Letters*, 9, 20130902.

Sydnor, Justin (2010). "(Over)insuring Modest Risks." *American Economic Journal: Applied Economics*, 2, 177–199.

Thakral, Neil and Linh To (2018). "Daily Labor Supply and Adaptive Reference Points." Available at http://scholar.harvard.edu/files/nthakral/files/ThakralTo_taxi.pdf.

Trautmann, Stefan T. and Gijs van de Kuilen (2018). "Higher Order Risk Attitudes: A Review of Experimental Evidence." *European Economic Review*, 103, 108–124.

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

Zeelenberg, Marcel (1999). "Anticipated Regret, Expected Feedback and Behavioral Decision Making." *Journal of Behavioral Decision Making*, 12, 93–106.

Supplementary Data

Supplementary data are available at *JEEA* online.