Salience and Skewness Preferences

Markus Dertwinkel-Kalt[1] and Mats Köster[2]

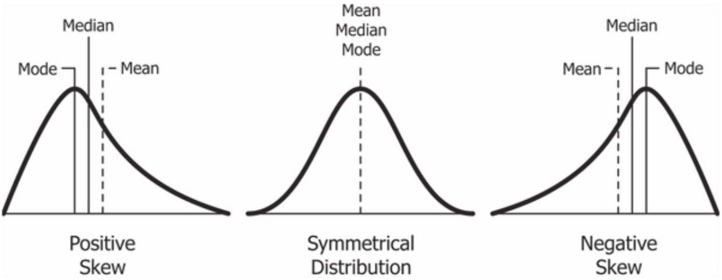
[1]Frankfurt School of Finance & Management

[2]Düsseldorf Institute for Competition Economics (DICE)

February 2019

The skewness of a probability distribution

Skewness typically refers to the central standardized third moment.



- Right-skewed = positively skewed: tail on the right side of the probability distribution is long $\rightarrow$ "large pos. payoff with a small probability."
- Left-skewed = negatively skewed: tail on the left side of the probability distribution is long $\rightarrow$ "large neg. payoff with a small probability."

Research questions and an overview of our results

Conventional wisdom: Most people prefer risks with a higher expected value and/or a lower variance. This can be overturned by **skewness preferences**:

people like right-skewed, but dislike left-skewed risks.

We argue that skewness preferences, typically attributed to CPT, are more naturally accommodated by salience theory (Bordalo et al. 2012, QJE):

1) How do risk attitudes depend on skewness according to salience theory?
   - Salience predicts a preference for right- & aversion toward left-skewed risks. Besides theoretical predictions, we further provide experimental support.

2) Does salience yield a preference for skewness after controlling for variance?
   - Yes, although *relative* rather than *absolute* skewness matters.
   - Relative Skewness: L_x is skewed relative to L_y if $L_x - L_y$ is right-skewed.
   - In a second lab experiment we manipulate relative skewness via the lotteries' correlation, which allows us to disentangle salience and CPT.

Focusing Illusion: contrasts attract attention

Central assumption: the **contrast effect** (“contrasts attract attention”).

- Dimensions along which the alternative options differ a lot attract a great deal of attention (e.g. Schkade and Kahneman 1998, PsyScience).
- Choice under risk: states with a large difference in attainable outcomes attract much attention and the corresponding probabilities are inflated.
- Also central role in Tversky (1969, PsyRev), Loomes and Sugden (1987, JET), Rubinstein (1988, JET), or Kőszegi and Szeidl (2013, QJE).
- Supportive lab evidence: e.g. Dertwinkel-Kalt and Köster (2017, JEBO) or Frydman and Mormann (2018, WP).
- Evidence from other domains: e.g. Hastings and Shapiro (2013, QJE) or Dertwinkel-Kalt et al. (2017, JEEA).

A preference for right-skewed risks & an aversion toward left-skewed risks

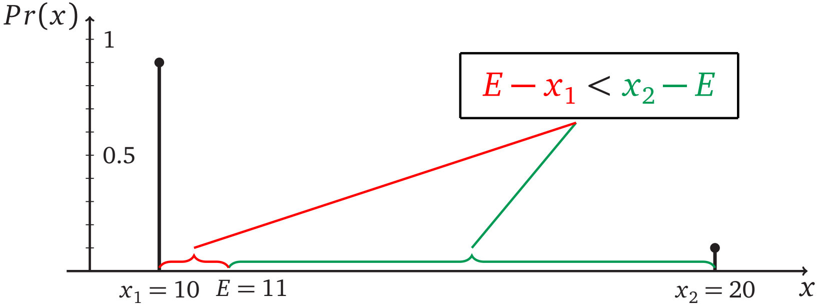
- Observation 1: People buy insurance against left-skewed risks.
  e.g. Sydnor (2010, AEJ) or Barseghyan et al. (2013, AER)
- Observation 2: People participate in right-skewed lottery games.
  e.g. Golec and Tamarkin (1998, JPE) or Garrett and Sobel (1999, EL)
- Observation 3: On asset markets, positive skewness is priced.
  e.g. Bali et al. (2011, JFE) or Conrad et al. (2013, JF)
- Observation 4: Workers accept a lower expected wage if the distribution of wages in a cluster (i.e., education-occupation combination) is right-skewed.
  e.g. Hartog and Vijverberg (2007, LE) or Grove et al. (2018, WP)
- Observation 5: Laboratory subjects prefer right-skewed over left-skewed risks with the same expected value and variance.
  e.g. Ebert and Wiesen (2011, MS) or Ebert (2015, JEBO)

A satisfying explanation for skewness preferences is missing

- While EUT might explain Observation 5 (e.g., Menezes et al. 1980, AER), it cannot explain why otherwise risk-averse people participate in unfair, but sufficiently right-skewed lottery games.
- CPT (Tversky and Kahneman 1992, JRU) assumes that probabilities of extreme events are overweighted, and may account for all observations.
- **But:** CPT predicts that only a lottery's absolute and **not** its relative skewness matters, which is inconsistent with our experimental findings.

The contrast effect and skewness preferences

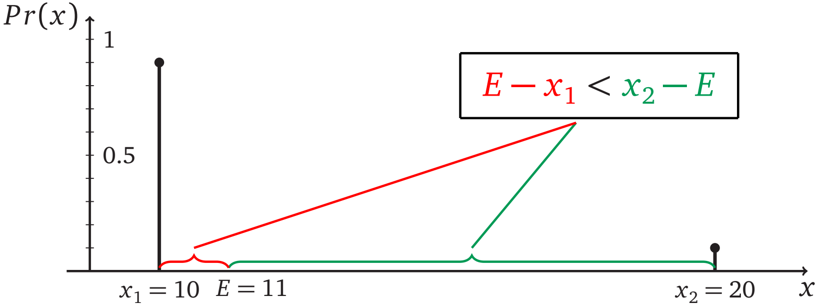
Consider the choice between a right-skewed binary lottery and a safe option E.



States of the world: (x_1, E) and (x_2, E)

The contrast effect and skewness preferences

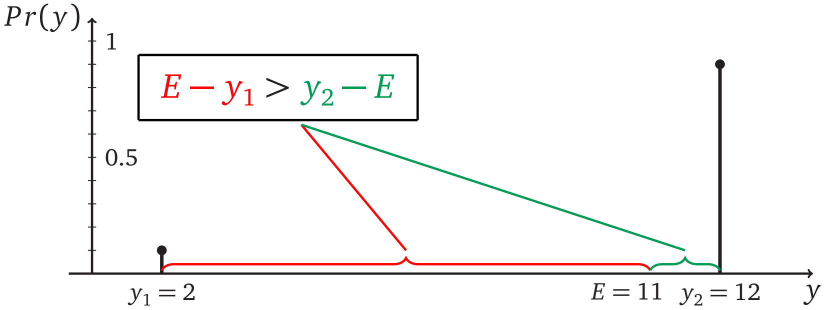
Consider the choice between a right-skewed binary lottery and a safe option E.



States of the world: (x_1, E) and (x_2, E) $\rightarrow$ (x_2, E) attracts more attention.

The contrast effect and skewness preferences

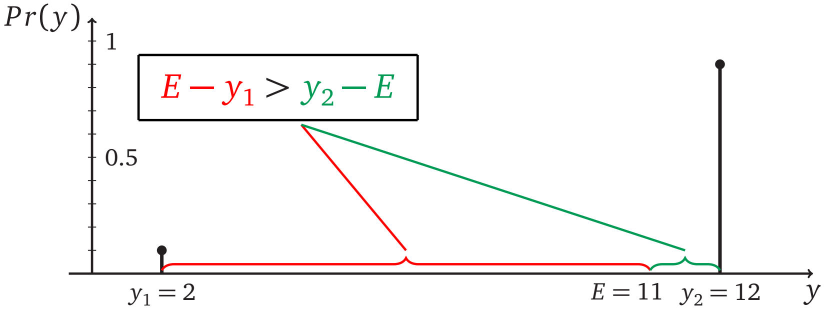
Consider the choice between a left-skewed binary lottery and a safe option E.



States of the world: (y_1, E) and (y_2, E)

The contrast effect and skewness preferences

Consider the choice between a left-skewed binary lottery and a safe option E.



States of the world: (y_1, E) and (y_2, E) $\rightarrow$ (y_1, E) attracts more attention.

The contrast effect and skewness preferences

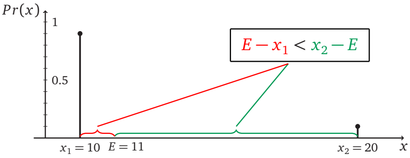


Figure: *Right-skewed vs. expected value.*

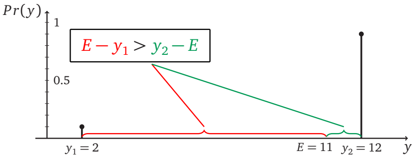


Figure: *Left-skewed vs. expected value.*

Note: Both lotteries have the same expected value and variance; i.e., both are "equally risky."

Salience model (Bordalo et al., 2012; henceforth: BGS)

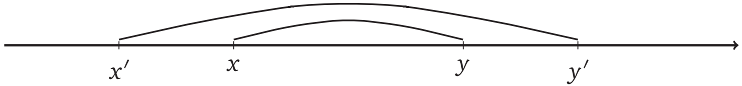
- An agent with linear value function chooses between lotteries L_x and L_y.
- The lotteries' joint distribution F determines the state space $S \subseteq \mathbb{R}^2$.
- The weight assigned to each state $s \in S$ depends on this state's salience.
- Salience is assessed via a symmetric and bounded function $\sigma : \mathbb{R}^2 \to \mathbb{R}_+$ that satisfies the **contrast effect** and the level effect (cont'd next slide).
- A salient thinker's decision utility from L_x in $\mathcal{C} := \{L_x, L_y\}$ is given by

$$U^s(L_x|\mathcal{C}) = \underbrace{\frac{1}{\int_{\mathbb{R}^2} \sigma(v,w)\ dF(v,w)}}_{\text{normalization factor}} \int_{\mathbb{R}^2} x \cdot \sigma(x,y)\ dF(x,y).$$

Fundamentals of salience theory

Contrast effect: differences attract a decision maker's attention.

$$\sigma(x', y') > \sigma(x, y)$$



Level effect: a given contrast is less salient at a higher outcome level.

$$\sigma(x, y) > \sigma(x+\epsilon, y+\epsilon)$$

Correlation determines the state space and a salient thinker's valuation of lotteries

If $L_x = (x_1, p; x_2, 1-p)$ and $L_y = (y_1, q; y_2, 1-q)$, the state space satisfies

$$S \subseteq \{(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)\},$$

whereby the exact number of states depends on the correlation structure:

- independence $\Rightarrow$ four states of the world;
- imperfect correlation $\Rightarrow$ three or four states;
- perfect correlation $\Rightarrow$ two states of the world.

$\longrightarrow$ Correlation affects salience of outcomes and thus a salient thinker's valuation.

Binary risks allow us to assess how skewness affects risk attitudes

For most classes of lotteries, it is **hard (or impossible)** to study skewness effects, as there exist several measures of skewness that are not equivalent in general.

But: for a binary lottery $L = (x_1, p; x_2, 1-p)$ with outcomes $x_1 < x_2$, skewness is unambigously defined by the lottery's third standardized central moment

$$S = \frac{2p-1}{\sqrt{p(1-p)}}.$$

Also, any binary risk is uniquely defined by its expected value E, its variance V, and its skewness S (Ebert 2015, JEBO), so that we can fix expected value and variance, and vary only the skewness of $L = L(E, V, S)$, which has parameters:

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

First contribution: Salience predicts skewness-dependent risk attitudes

Let $\mathcal{C} = \{L(E, V, S), E\}$.

Proposition 1

For any expected value E and variance V, there exists some $\hat{S} = \hat{S}(E, V) \in \mathbb{R}$ such that a salient thinker chooses the lottery if and only if $S > \hat{S}$.

Suppose that the salience function satisfies a decreasing level effect (most of the salience functions that we are aware of satisfy this property). Definition

Proposition 2

For any lottery $L(E, V, \hat{S}(E, V))$ with positive payoffs and any $\epsilon > 0$, we obtain

$$0 < \hat{S}(E + \epsilon, V) < \hat{S}(E, V).$$

Lotteries to test for skewness-dependent risk attitudes

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

Table: *Lotteries to test for Propositions 1 and 2; all have the same variance* $V = 225$.

Lotteries to test for skewness-dependent risk attitudes

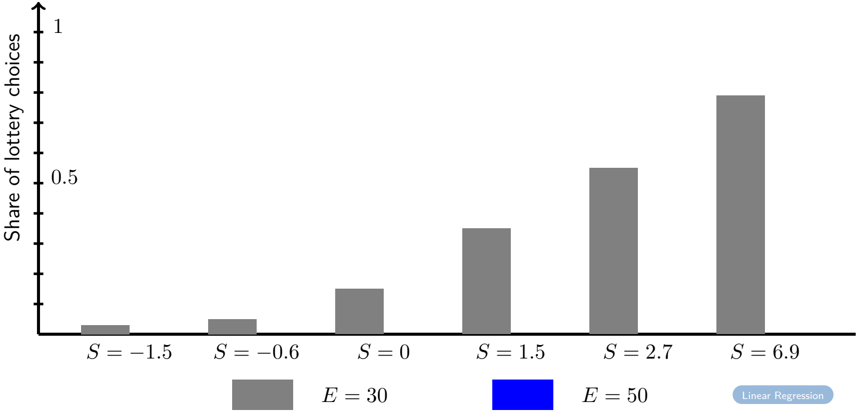
Lottery	Exp. Value	Skewness
(37.5, 80%; 0, 20%)	30	-1.5
(41.25, 64%; 10, 36%)	30	-0.6
(45, 50%; 15, 50%)	30	0
(60, 20%; 22.5, 80%)	30	1.5
(75, 10%; 25, 90%)	30	2.7
(135, 2%; 27.85, 98%)	30	6.9
(57.5, 80%; 20, 20%)	50	-1.5
(61.25, 64%; 30, 36%)	50	-0.6
(65, 50%; 35, 50%)	50	0
(80, 20%; 42.5, 80%)	50	1.5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

Table: *Lotteries to test for Propositions 1 and 2; all have the same variance* $V = 225$.

Experimental implementation

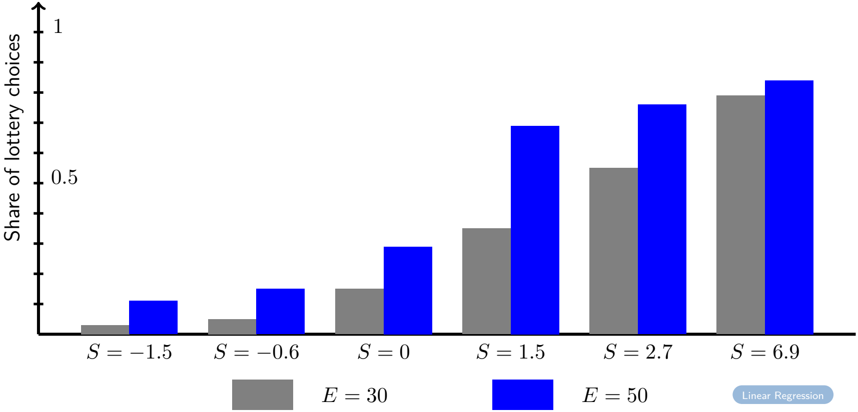
- $n = 62$ in 3 sessions in Jan 2018,
- 2 ECU = 1 Euro,
- one decision randomly picked and paid,
- random order of tasks.

Experimental results



Note: The figure illustrates the share of lottery choices for a low and a high expected value. The skewness values are presented in ascending order, but not in a proper scale.

Experimental results



Linear Regression

Note: The figure illustrates the share of lottery choices for a low and a high expected value. The skewness values are presented in ascending order, but not in a proper scale.

Mao pairs allow us to disentangle a preference for variance and skewness

Definition 1 (Mao pair)

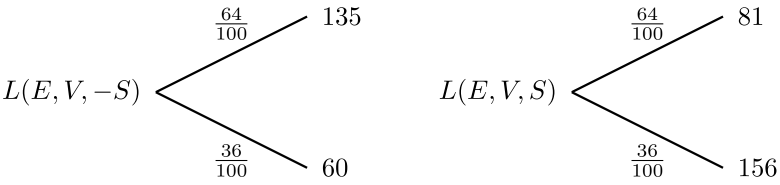
Let $S \in (0, \infty)$*. The lotteries* $L(E, V, S)$ *and* $L(E, V, -S)$ *denote a* Mao pair.

Mao pairs allow us to disentangle a preference for variance and skewness

Definition 1 (Mao pair)

Let $S \in (0, \infty)$*. The lotteries* $L(E, V, S)$ *and* $L(E, V, -S)$ *denote a* Mao pair.

Example: Let $E = 108$, $V = 1296$, and $S = 0.6$.

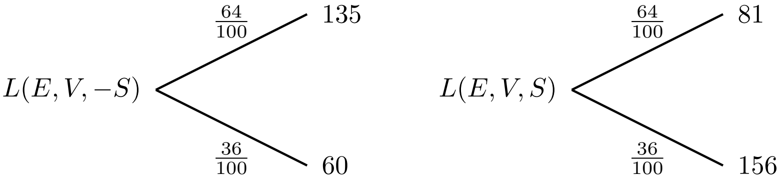


Mao pairs allow us to disentangle a preference for variance and skewness

Definition 1 (Mao pair)

Let $S \in (0, \infty)$*. The lotteries* $L(E, V, S)$ *and* $L(E, V, -S)$ *denote a* Mao pair.

Example: Let $E = 108$, $V = 1296$, and $S = 0.6$.



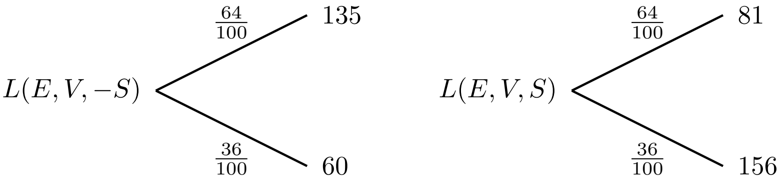
Here, the state space depends on the correlation structure.

Mao pairs allow us to disentangle a preference for variance and skewness

Definition 1 (Mao pair)

Let $S \in (0, \infty)$. *The lotteries* $L(E, V, S)$ *and* $L(E, V, -S)$ *denote a* Mao pair.

Example: Let $E = 108$, $V = 1296$, and $S = 0.6$.



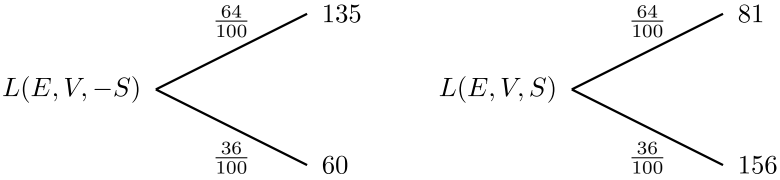
- Perfectly negative correlation: $(135, 81)$ and $(60, 156)$.

Mao pairs allow us to disentangle a preference for variance and skewness

Definition 1 (Mao pair)

Let $S \in (0, \infty)$. *The lotteries* $L(E, V, S)$ *and* $L(E, V, -S)$ *denote a* Mao pair.

Example: Let $E = 108$, $V = 1296$, and $S = 0.6$.



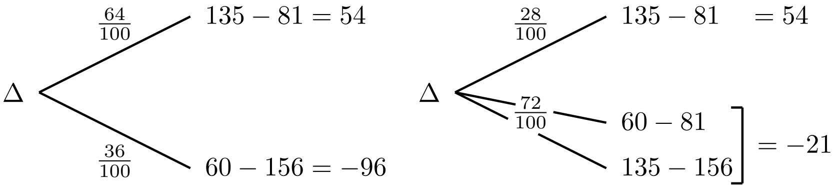
- Perfectly negative correlation: $(135, 81)$ and $(60, 156)$.
- Maximal positive correlation: $(135, 81)$, $(60, 81)$, and $(135, 156)$.

Relative skewness varies with the correlation structure

Definition 2 (Relative Skewness)

A lottery L_x is skewed relative to L_y if and only if $\Delta = L_x - L_y$ is right-skewed.

Example: Let $\Delta = L(E, V, -S) - L(E, V, S)$, and $E = 108$, $V = 1296$, $S = 0.6$.



perfectly negative correlation

maximal positive correlation

Relative skewness varies with the correlation structure – cont'd

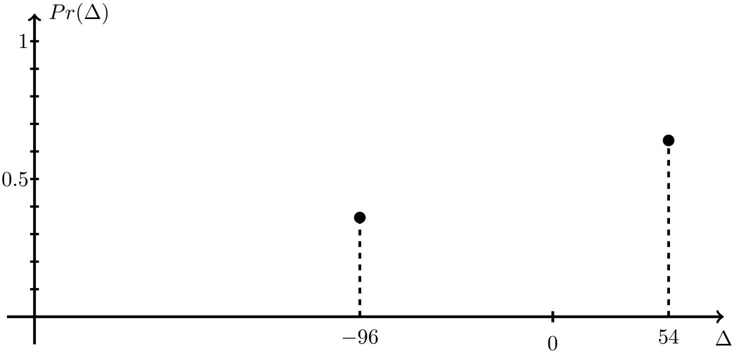


Figure: *Distribution of Δ under perfectly negative correlation.*

Relative skewness varies with the correlation structure – cont'd

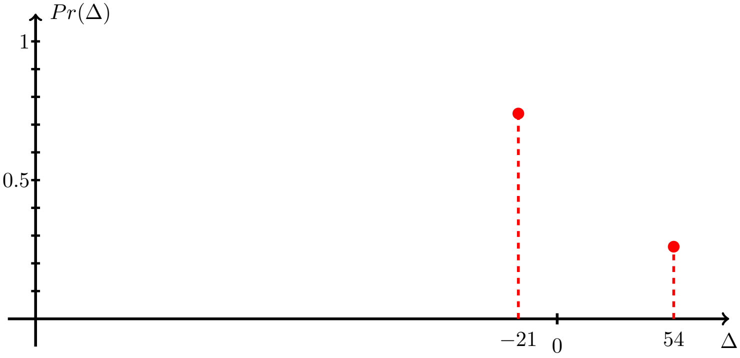


Figure: *Distribution of Δ under maximal positive correlation.*

Relative skewness varies with the correlation structure – cont'd

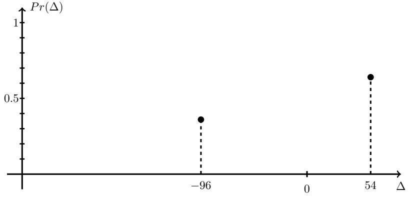


Figure: *Perfectly negative correlation.*

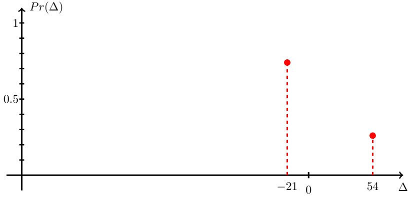


Figure: *Maximal positive correlation.*

$\rightarrow$ Δ is left-skewed under negative and right-skewed under positive correlation.

Not only correlation, but also absolute skewness determines relative skewness

Consider a *more skewed* Mao pair—i.e., $S = 2.7$ instead of $S = 0.6$—with the same expected value and the same variance.

Figure: *Perfectly negative correlation.*

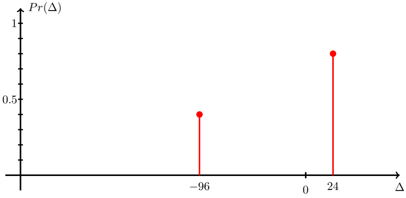


Figure: *Maximal positive correlation.*

$\rightarrow$ Δ is left-skewed both under negative and under positive correlation.

More generally: Δ is always left-skewed under perfectly negative correlation and becomes right-skewed under maximal positive correlation if and only if $S < 1.15$.

Second contribution: Salient thinkers like relative rather than absolute skewness

Proposition 3

For any Mao pair, there exists some $\check{S} > 0$ such that the following holds:

(a) *Under the perfectly negative correlation, a salient thinker always prefers $L(E, V, S)$ over $L(E, V, -S)$.*

(b) *Under the maximal positive correlation, a salient thinker prefers $L(E, V, S)$ over $L(E, V, -S)$ if and only if $S \geq \check{S}$.*

$\longrightarrow$ Salience predicts a (larger) shift towards the left-skewed lottery for small S.

$\longrightarrow$ This prediction is consistent with a **preference for relative skewness**.

The experimental Mao pairs

Left-skewed Lottery	Right-skewed Lottery	Exp. Value	Var.	Skewness
(120, 90%; 0, 10%)	(96, 90%; 216, 10%)	108	1296	$\pm$ 2.7
(135, 64%; 60, 36%)	(81, 64%; 156, 36%)	108	1296	$\pm$ 0.6
(40, 90%; 0, 10%)	(32, 90%; 72, 10%)	36	144	$\pm$ 2.7
(45, 64%; 20, 36%)	(27, 64%; 52, 36%)	36	144	$\pm$ 0.6
(80, 90%; 0, 10%)	(64, 90%; 144, 10%)	72	576	$\pm$ 2.7
(90, 64%; 40, 36%)	(54, 64%; 104, 36%)	72	576	$\pm$ 0.6

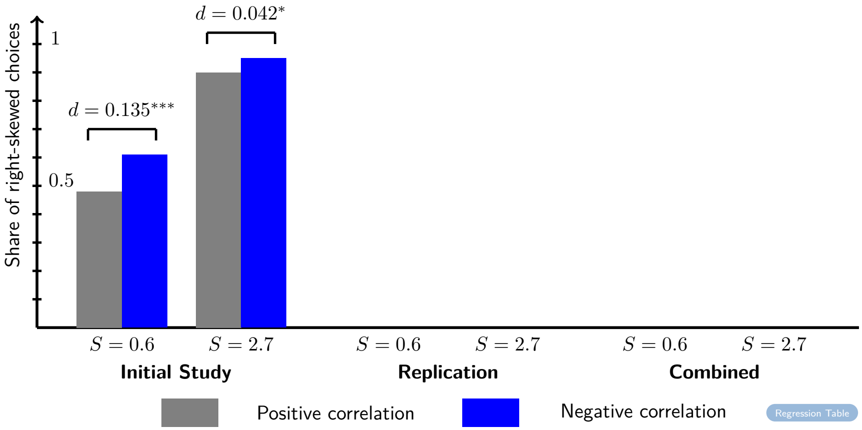
Table: *Mao pairs to study the effect of correlation on choice under risk.*

Decision Screen

Experimental implementation

- $n = 79$ in 3 sessions in Feb and Mar 2018,
- replication study in Nov 2018 with $n = 113$,
- 4 ECU $=$ 1 Euro,
- one decision randomly picked and paid,
- random order of tasks.

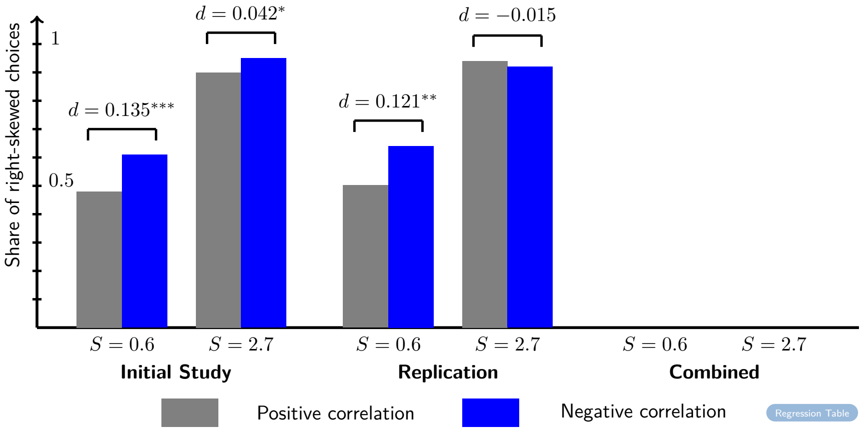
Experimental results



Regression Table

Note: The figure illustrates the share of choices in favor of the right-skewed lottery under positive and negative correlation. We also report the results of paired t-tests with standard errors being clustered at the subject level. Significance level: *: 10%, **: 5%, ***: 1%.

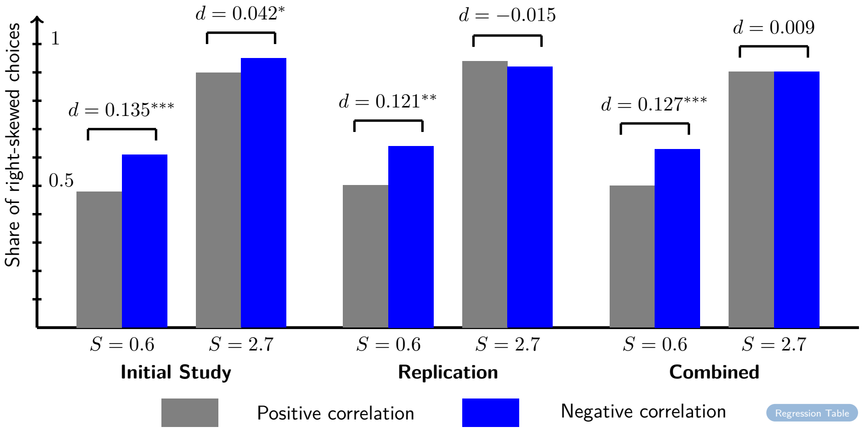
Experimental results



Regression Table

Note: The figure illustrates the share of choices in favor of the right-skewed lottery under positive and negative correlation. We also report the results of paired t-tests with standard errors being clustered at the subject level. Significance level: *: 10%, **: 5%, ***: 1%.

Experimental results



Regression Table

Note: The figure illustrates the share of choices in favor of the right-skewed lottery under positive and negative correlation. We also report the results of paired t-tests with standard errors being clustered at the subject level. Significance level: *: 10%, **: 5%, ***: 1%.

Other models based on the contrast effect also predict skewness preferences

A Model of Focusing (Kőszegi and Szeidl 2013, QJE):

- The focusing function—the pendant to the salience function—satisfies the contrast, but not the level effect.
- But Kőszegi and Szeidl assume a non-linear value function, such that, for binary choices, we can closely align focusing and salience theory.

Generalized Regret Theory (Loomes and Sugden 1987, JET):

- Lanzani (2018, WP) shows that for binary choices salience is a special case of generalized regret theory. But underlying psychology is very different.
- Zeelenberg (1999, JBDM) finds that regret affects decisions only if subjects know that they will receive feedback on the counterfactal outcome.
- Our results impose restrictions on regret model: rejoice has to matter. Details
- And, with more than two options, we can really tell the theories apart:
  - Dertwinkel-Kalt and Köster (2017, JEBO): dominated decoys.
  - Frydman and Mormann (2018, WP): phantom decoys.

Conclusion

- Applying salience theory to simple choice problems, we have unraveled the *contrast effect* as a plausible driver of skewness preferences.
- Skewness preferences are a robust observation, not only in humans, but also in animals (Strait and Hayden 2013, Bio Letters; Genest et al. 2016, PNAS).
- We further provide evidence suggesting that not only absolute but also rel. skewness matters, which is consistent with salience but not with CPT.
- Skewness preferences and, in particular, a preference for relative skewness may help us to better understand other phenomena like the Allais paradoxes.

Common-consequence Allais

Thank you for your attention!

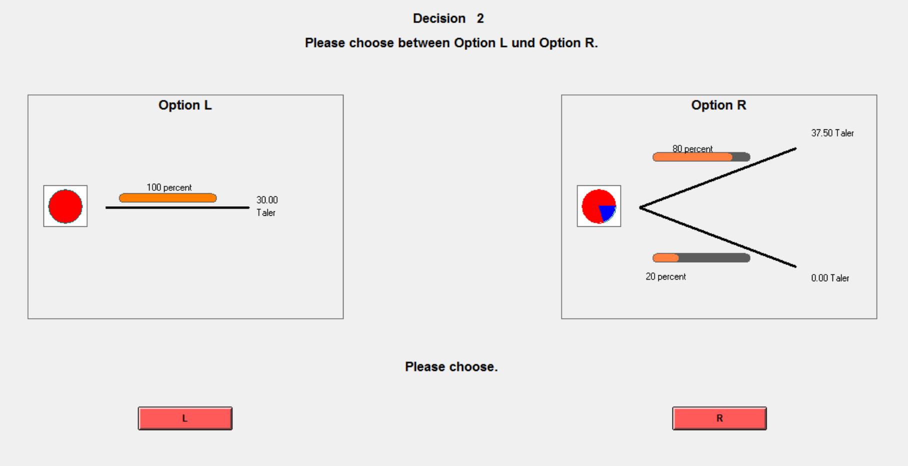
Definition of the decreasing level effect

Definition 3 (Decreasing Level Effect)

Suppose that $x, y, z \in \mathbb{R}$ with $x + y, x + z \geq 0$. For a given salience function σ, let $\varepsilon_\sigma(x, y, z) := -\frac{\frac{d}{dx}\sigma(x+y,x+z)}{\sigma(x+y,x+z)}$. The salience function σ satisfies a decreasing level effect *if and only if $\varepsilon_\sigma(x, y, z)$ and $\varepsilon_\sigma(-x, -y, -z)$ decrease in y and z.*

Back

Decision screen in the first experiment: low expected value



Back

Main regression for the first experiment

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

Table: *OLS with clustered standard errors. Significance: *: 10%, **: 5%, ***: 1%.*

Unit of Observation: choice between a lottery and its expected value.

Dependent Variable: $Y = 1$ if subject chooses the lottery and $Y = 0$ otherwise.

Independent Variables: High Expected Value $= 1$ if $E = 50$ and High Expected Value $= 0$ otherwise. Skewness is a continuous variable.

Back

Decision screen in the second experiment: maximal positive correlation

Decision 1

Please choose between Option A und Option B.

	Fields 1-36	Fields 37-72	Fields 73-100
Option A	**90**	**40**	**90**
Option B	**104**	**54**	**54**

A

B

Main regression for the second experiment

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

Table: *OLS with clustered standard errors. Significance: *: 10%, **: 5%, ***: 1%.*

Unit of Observation: the pair of choices corresponding to the same Mao pair.

Dependent Variable: $Y = 1$ if subject switches from right-skewed under perfectly negative to left-skewed under maximal positive correlation, $Y = -1$ if the subject switches in the opposite direction, and $Y = 0$ if the subject does not switch.

Independent Variable: Skewed = 1 if $S = 2.7$ and Skewed = 0 otherwise.

Back

Original Regret Theory (Loomes and Sugden 1982, EJ)

Two lotteries L_x and L_y. State $s_{ij} = (x_i, y_j) \in S$ occurs with prob. $\pi_{ij} > 0$.

There exists an increasing function $Q : \mathbb{R} \to \mathbb{R}$ with (i) $Q(z) = -Q(-z)$ and (ii) $Q : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ being convex, and an increasing function $c : \mathbb{R} \to \mathbb{R}$ such that

$$L_x \succeq L_y \quad \Longleftrightarrow \quad \sum_{s_{ij} \in S} \pi_{ij} Q(c(x_i) - c(y_j)) \geq 0.$$

Notably, this does **not** imply that the decision maker experiences rejoice.

Let the agent's only objective be to **minimize regret** or, formally, maximize

$$U^R(L_x | \{L_x, L_y\}) = \sum_{s_{ij} \in S} \pi_{ij} \min\{Q(c(x_i) - c(y_j)), 0\}.$$

This objective is consistent with the above definition, as

$$U^R(L_x | \{L_x, L_y\}) - U^R(L_y | \{L_x, L_y\}) = \sum_{s_{ij} \in S} \pi_{ij} Q(c(x_i) - c(y_j)).$$

Minimizing regret cannot explain the findings in Experiment 1

By choosing the safe option, subjects can completely **rule out** any regret:

"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."

$\longrightarrow$ If the safe option is chosen, there is not even a counterfactual outcome.

In summary, assuming that subjects experience rejoice, is **necessary** for regret to explain our results. Our results further impose restrictions on functional forms:

$$\text{Propositions 1 \& 2} \implies c''(\cdot) < 0 \text{ and } \frac{c'(E) - c'(x_2)}{c'(x_1) - c'(E)} < \frac{H(c(x_2) - c(E))}{H(c(E) - c(x_1))},$$

where $H(z) := Q(z)/Q'(z) \longrightarrow$ more curvature in Q implies more curvature c.

Back

Common-consequence Allais paradox depends on correlation (Frydman and Mormann 2018, WP)

$L^1(z) = (25, 0.33; 0, 0.01; z, 0.66)$ and $L^2(z) = (24, 0.34; z, 0.66)$ for $z \in \{0, 24\}$.

$z = 24$: aversion toward left-skewed risks $\longrightarrow$ majority chooses $L^2(24)$.

$z = 0$: choice depends on the correlation structure, as follows:

- independence $\longrightarrow$ 51% choose $L^1(0)$;
- intermediate correlation $\longrightarrow$ 41% choose $L^1(0)$;
- maximal correlation $\longrightarrow$ 20% choose $L^1(0)$.

Notably, the relative skewness of $L^2(0)$ increases as we move from

independence to intermediate correlation to maximal correlation.

$\longrightarrow$ Experimental results are consistent with a preference for relative skewness.

Back