Pre-Analysis Plan: Salience Effects in Portfolio Selection

Markus Dertwinkel-Kalt* Mats Köster†

January 2020

1 Aim of the Study

A large experimental literature on choices between two reduced (or simple) lotteries has documented two robust facts: First, if both lotteries are *symmetric*, then subjects are typically risk averse; that is, with symmetric risks a large majority of subjects chooses the option with less variance in outcomes (e.g., Kahneman and Tversky, 1979). Second, if at least one of the lotteries is *skewed*,[1] subjects often reveal a preference for positively skewed or right-skewed risks, which can even result in risk-seeking behavior; that is, a large majority of subjects chooses a sufficiently right-skewed risk even if it is the option with more variance in outcomes (e.g., Kahneman and Tversky, 1979; Dertwinkel-Kalt and Köster, forthcoming).

In this study, we ask how robust the observed patterns in these simple choices are to increasing the complexity of the problem along two dimensions: we increase the number of available options and/or replace the reduced lotteries by portfolios (i.e., convex combinations of reduced lotteries), where the latter means that the final outcomes are not stated, but have to be inferred. The experiment is designed to test the predictions of *salience theory of choice under risk* (Bordalo *et al.*, 2012) for such portfolio selection problems against those of naive decision rules like the $1/N$-heuristic (e.g., Benartzi and Thaler, 2001; Eyster and Weizsäcker, 2016), which predict that subjects *diversify naively* by investing equal shares into the available reduced lotteries.

2 An Application of Salience Theory to Portfolio Selection

Consider an agent who decides how much to invest in either of two non-negative random variables X_1 and X_2 (also called *assets*) with a joint cumulative distribution function F. The agent thus forms a portfolio $X(\alpha) := \alpha X_1 + (1-\alpha)X_2$, where any share $\alpha \in A \subseteq [0,1]$ is admissable; that is, the agent's choice set is given by $\mathcal{C} := \{X(\alpha) : \alpha \in A\}$. We study portfolio selection problems of a varying complexity, as captured by (1) the cardinality of $\mathcal{C}$, ranging from binary to continuous choice sets, and (2) the difficulty of determining the distribution of outcomes.

*Frankfurt School of Finance & Management, Adickesallee 32-34, 60332 Frankfurt, Germany. Email: m.dertwinkel-kalt@fs.de.

†HHU Düsseldorf (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany. Email: mats.koester@hhu.de.

[1] A lottery's skewness is typically measured by its third standardized central moment. A lottery is symmetric/right-skewed/left-skewed if its third standardized central moment is equal to/larger/smaller than zero.

We assume that the agent is a *salient thinker*, as introduced in Bordalo *et al.* (2012), who evaluates any given portfolio, $X(\alpha)$, relative to the *reference portfolio*

$$\overline{X}(\alpha) := \int_{A\setminus\{\alpha\}} X(\beta)\, dG_\alpha(\beta), \tag{1}$$

where $G_\alpha(\cdot)$ is the cumulative distribution function of the uniform distribution over the augmented choice set $\mathcal{C}_\alpha := \{X(\beta) : \beta \in A,\ \beta \neq \alpha\}$. For any binary choice set $\mathcal{C} = \{X(\alpha), X(\beta)\}$, the reference portfolio to $X(\alpha)$ is thus given by the alternative option $X(\beta)$ and vice versa.

According to salience theory of choice under risk (Bordalo *et al.*, 2012), a decision-maker evaluates a portfolio by assigning a subjective probability to each state of the world that depends on the state's objective probability, as given by F, and on how *salient* this state is for this particular portfolio (relative to the reference portfolio). Precisely, a state of the world is given by a realization $(x_1, x_2) \in \mathbb{R}_+^2$ of the assets' joint distribution, which corresponds, for given share $\alpha \in A$, to a tuple $(x(\alpha), \overline{x}(\alpha))$—where the reference point $\overline{x}(\alpha)$ is defined as in Eq. (1)—that then determines how salient this particular state is for this particular portfolio. More generally, the salience of any tuple $(x, y) \in \mathbb{R}_+^2$ is determined by a so-called *salience function* $\sigma : \mathbb{R}_+^2 \to \mathbb{R}_+$, as defined below, and a tuple $(x, y) \in \mathbb{R}_+^2$ is said to be the more salient the larger $\sigma(x, y)$ is.

Definition 1 (Salience Function)**.** *We say that a symmetric, bounded, and absolutely continuous function $\sigma : \mathbb{R}_+^2 \to \mathbb{R}_+$ is a* salience function *if and only if it satisfies the following two properties:*

1. Ordering. *Let $x \geq y$. Then, for any $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$,*

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

2. Diminishing sensitivity. *Let $x, y \geq 0$ with $x \neq y$. Then, for any $\epsilon > 0$,*

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

For certain setups, we have to impose more structure on the salience function (i.e., homogeneity of degree zero) to derive clear-cut predictions, as assumed by Bordalo *et al.* (2013, 2016). We will indicate explicitly any results that rely on this additional assumption:

Assumption 1 (Homogeneity of Degree Zero)**.** *For any $x, y \geq 0$ and $\lambda > 0$, $\sigma(x, y) = \sigma(\lambda x, \lambda y)$.*

We use the continuous variant of salience theory proposed by Dertwinkel-Kalt and Köster (forthcoming). A salient thinker evaluates outcomes via a linear value function, $u(x) = x$, and chooses from the choice set $\mathcal{C}$ as to maximize her the salience-weighted utility defined as follows.

Definition 2 (Salience-Weighted Utility)**.** *Fix a choice set $\mathcal{C}$ and let $\overline{x}(\alpha)$ be defined as in Eq.* (1)*. The* salience-weighted utility *of a portfolio $X(\alpha)$ evaluated in the choice set $\mathcal{C}$ equals*

$$U^s(X(\alpha)|\mathcal{C}) = \int_{\mathbb{R}_+^2} (\alpha x_1 + (1-\alpha)x_2) \cdot \frac{\sigma(\alpha x_1 + (1-\alpha)x_2, \overline{x}(\alpha))}{\int_{\mathbb{R}_+^2} \sigma(\alpha y_1 + (1-\alpha)y_2, \overline{y}(\alpha))\, dF(y_1, y_2)}\, dF(x_1, x_2),$$

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

3 Theoretical Analysis of Portfolio Selection Problems

3.1 Setup and Definitions

We assume that the two assets X_1 and X_2 are binary lotteries, which are uniquely characterized by their first three moments: expected value, variance, and skewness (see Ebert, 2015).

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

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

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

To identify the effect of skewness on behavior, we use a certain class of binary lotteries that was introduced by Mao (1970) and later formalized by Ebert and Wiesen (2011, Definition 2).

Definition 3. *Let $S \in \mathbb{R}_+$. The random variables $L(E, V, S)$ and $L(E, V, -S)$ denote a* Mao pair*.*

The two lotteries of a Mao pair have the exact same expected value and variance, but the opposite skewness. The joint distribution of two lotteries that form a Mao pair can be fully parameterized by their correlation coefficient

$$\rho = \rho\big(L(E, V, -S), L(E, V, S)\big) \in [-1, \overline{\rho}(S)] \quad \text{with} \quad \overline{\rho}(S) := \frac{\sqrt{4+S^2} - S}{\sqrt{4+S^2} + S}.$$

Let $p = p(-S) \in (0, \frac{1}{2})$ be the probability of the left-skewed lottery's smaller payoff, which is, by construction, identical to the probability of the right-skewed lottery's larger payoff. For a given correlation coefficient ρ, Table 1 depicts the joint distribution of a Mao pair:

	$(1-p)p(1+\rho)$	$p^2 - (1-p)p\rho$	$(1-p)^2 - (1-p)p\rho$	$(1-p)p(1+\rho)$
$L(E, V, -S)$	$E - \sqrt{\frac{V(1-p)}{p}}$	$E - \sqrt{\frac{V(1-p)}{p}}$	$E + \sqrt{\frac{Vp}{(1-p)}}$	$E + \sqrt{\frac{Vp}{(1-p)}}$
$L(E, V, S)$	$E - \sqrt{\frac{Vp}{(1-p)}}$	$E + \sqrt{\frac{V(1-p)}{p}}$	$E - \sqrt{\frac{Vp}{(1-p)}}$	$E + \sqrt{\frac{V(1-p)}{p}}$

Table 1: *Joint distribution of the lotteries of a Mao pair.*

We will test the predictions of salience theory relative to the benchmarks of choosing a *diversified* portfolio in the sense of Markowitz (1952) or a *naively diversified* portfolio, as it would be predicted by naive decision rules such as the $1/N$-heuristic (e.g., Benartzi and Thaler, 2001).

Definition 4 (Diversification)**.** *We say that a portfolio $X(\alpha)$ is* diversified *if and only if there does not exist any other portfolio $X(\alpha')$ with $\mathbb{E}[X(\alpha')] \geq \mathbb{E}[X(\alpha)]$ and $Var(X(\alpha')) < Var(X(\alpha))$. Any portfolio $X(\alpha)$ that is not diversified is said to be* under-diversified*.*

Definition 5 (Naive Diversification)**.** *A portfolio $X(\alpha)$ is* naively diversified *if and only $\alpha = \frac{1}{2}$.*

3.2 Portfolio Selection with Skewed Assets

Binary choice sets. In this section, we will rely on assets that form a Mao pair, because the resulting portfolios have nice properties, summarized in the following lemma. In particular, irrespective of the correlation structure, the diversified portfolio coincides with the naively diversified portfolio, so that here under-diversification cannot be explained by the $1/N$-heuristic.

Lemma 2. *Let $X_1 = L(E, V, S)$ and $X_2 = L(E, V, -S)$. Then, the variance of the portfolio $X(\alpha)$ increases in the distance $|\alpha - \frac{1}{2}|$ and is therefore minimized at $\alpha = \frac{1}{2}$.*

If, in addition, the assets are perfectly negatively correlated, then the following two statements hold:

(*a*) *For any $\alpha \in [0, 1]$, we have $X(\alpha) = L(E, (2\alpha - 1)^2 V, \mathrm{sgn}(2\alpha - 1)S)$.*

(*b*) *For any $\alpha \neq \frac{1}{2}$, the portfolios $X(\alpha)$ and $X(1 - \alpha)$ constitute a Mao pair.*

We analyze a salient thinker's binary choice between the diversified portfolio, $X(\frac{1}{2})$, and some under-diversified portfolio, $X(\alpha)$ with $\alpha \neq \frac{1}{2}$, under different correlation structures.

Proposition 1. *Let $A = \{\alpha, \frac{1}{2}\}$ with $\alpha \neq \frac{1}{2}$. Under the perfectly negative correlation, a salient thinker chooses $X(\alpha)$ over $X(\frac{1}{2})$ only if $\alpha > \frac{1}{2}$ holds; i.e., only if the under-diversified portfolio is right-skewed. For any $\alpha > \frac{1}{2}$, there exists some $\hat{S} \in \mathbb{R}$, such that for any $S > \hat{S}$, she chooses $X(\alpha)$ over $X(\frac{1}{2})$.*

Under the perfectly negative correlation, the diversified portfolio eliminates all the variance in outcomes and pays the portfolio's expected value with certainty. As a consequence, by Lemma 2 (a), the problem boils down to the question whether a salient thinker chooses the non-negative binary lottery $L(E, (2\alpha-1)^2 V, \mathrm{sgn}(2\alpha-1)S)$ over its expected value or not. Here, diminishing sensitivity implies that a salient thinker will never choose a left-skewed portfolio with $\alpha < \frac{1}{2}$ (see Dertwinkel-Kalt and Köster, forthcoming, Corollary 1). If the Mao pair is sufficiently skewed, however, a salient thinker chooses a right-skewed, but under-diversified portfolio with $\alpha > \frac{1}{2}$ (see Proposition 3 in Dertwinkel-Kalt and Köster, forthcoming). To derive predictions for other correlation structures, in particular, for the case of the maximal positive correlation, we need to impose more structure on the salience function (e.g., Assumption 1).

Assuming homogeneity of degree zero pins down the trade-off between ordering and diminishing sensitivity; in particular, it imposes that the latter is sizeable. As a consequence, under the perfectly negative correlation, a salient thinker does not choose certain right-skewed portfolios over the diversified one. Under the maximal positive correlation, we observe that for sufficiently skewed Mao pairs a salient thinker chooses an under-diversified portfolio if and only if it is right-skewed, while for sufficiently symmetric Mao pairs the opposite is true.

Corollary 1. *Let $A = \{\alpha, \frac{1}{2}\}$ with $\alpha \neq \frac{1}{2}$ and $p = p(-S)$ as in Lemma 2. Suppose Assumption 1 hold.*

(*a*) *Under the perfectly negative correlation, a salient thinker chooses $X(\alpha)$ over $X(\frac{1}{2})$ if and only if*

$$\alpha \in \left(\frac{1}{2}, \frac{1}{2} + \frac{ES}{\sqrt{4V}} \right).$$

In particular, if $S \geq \sqrt{V}/E$, a salient thinker chooses $X(\alpha)$ over $X(\frac{1}{2})$ if and only if $\alpha > \frac{1}{2}$.

(b) *Under the maximal positive correlation, a salient thinker chooses $X(\alpha)$ over $X(\frac{1}{2})$ for any*

$$\alpha \in \left[\frac{E}{\sqrt{V}}\sqrt{\frac{1-p}{p}}\left(\frac{1-4p}{2-4p}\right)+\frac{1-p}{1-2p}, \frac{1}{2}\right) \cup \left(\frac{1}{2}, \frac{E}{\sqrt{V}}\sqrt{\frac{1-p}{p}}\left(\frac{1-4p}{2-4p}\right)-\frac{p}{1-2p}\right],$$

but she chooses $X(\frac{1}{2})$ over $X(\alpha)$ for any

$$\alpha \leq \min\left\{\frac{1}{2}, \frac{E}{\sqrt{V}}\sqrt{\frac{1-p}{p}}\left(\frac{1-4p}{2-4p}\right)-\frac{p}{1-2p}\right\} \text{ and } \alpha \geq \max\left\{\frac{1}{2}, \frac{E}{\sqrt{V}}\sqrt{\frac{1-p}{p}}\left(\frac{1-4p}{2-4p}\right)+\frac{1-p}{1-2p}\right\}.$$

In particular, if $2S-\sqrt{S^2+4} \leq -2\sqrt{V}/E$, then a salient thinker chooses $X(\alpha)$ over $X(\frac{1}{2})$ if and only if $\alpha < \frac{1}{2}$. If $2S-\sqrt{S^2+4} \geq 2\sqrt{V}/E$, she chooses $X(\alpha)$ over $X(\frac{1}{2})$ if and only if $\alpha > \frac{1}{2}$.

The predictions derived in Corollary 1 are illustrated in Figure 1. The case of the perfectly negative correlation (i.e., Part (a) of Corollary 1) is depicted in the left panel, while the case of the maximal positive correlation (i.e., Part (b) of Corollary 1) is depicted in the right panel.

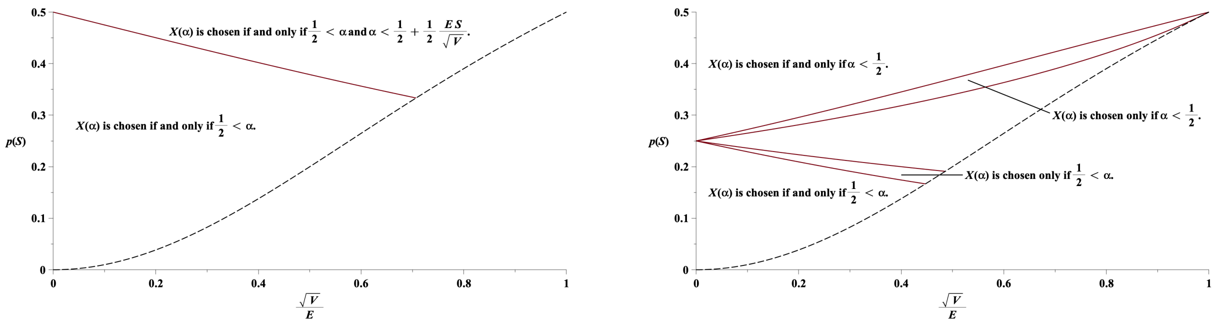

Figure 1: *Illustration of Corollary 1.*

Continuous choice sets. Next, we increase the complexity of the portfolio selection problem by studying the choice from a continuous set of options. The assets still constitute a Mao pair.

Proposition 2. *Let $A = [0,1]$. Under the perfectly negative correlation, a salient thinker selects a portfolio with $\alpha \geq \frac{1}{2}$. Moreover, if a salient thinker selects, for a fixed expected value E and variance V, a portfolio with $\alpha > \frac{1}{2}$ for some skewness level S', then she does so for any skewness level $S > S'$.*

This result again follows from the fact that a salient thinker exhibits a preference for right- and an aversion toward left-skewed binary risks, and it extends Proposition 1 to the domain of continuous choice sets. As before, if we impose more structure on the salience function by assuming homogeneity of degree zero, we obtain a more precise prediction.

Corollary 2. *Let $A = [0,1]$ and suppose that Assumption 1 holds. Under the perfectly negative correlation, a salient thinker selects an under-diversified portfolio with*

$$\alpha \in \left(\frac{1}{2}, \frac{1}{2}+\frac{ES}{\sqrt{4V}}\right)$$

It is important to notice that a salient thinker does not simply trade-off the skewness of portfolio returns against the variance in portfolio returns, as it is typically assumed in the empirical finance literature (e.g., Mitton and Vorkink, 2007). Under the perfectly negative correlation, the portfolio's variance monotonically increases in α on $(\frac{1}{2}, 1]$, but the portfolio's skewness is independent of α (see Lemma 2). Since $\alpha = \frac{1}{2}$ is a feasible choice, but not optimal according to salience theory, a salient thinker does not minimize the variance in portfolio returns for a given level of skewness. In other words, the positive skewness that is associated with choosing a portfolio with $\alpha > \frac{1}{2}$ renders (some) variance attractive to a salient thinker. Thus, our salience model makes predictions that are qualitatively different from those of the models that typically used to (poorly) explain skewness preferences in portfolio selection.

Finally, we consider the case of the maximal positive correlation, where we have to impose Assumption 1 in order to predict a salient thinker's behavior. As it is illustrated in Figure 2, a salient thinker never selects a left-skewed portfolio for sufficiently skewed Mao pairs, while she never selects a right-skewed portfolio for sufficiently symmetric Mao pairs.

Proposition 3. *Let $A = [0, 1]$ and suppose that Assumption 1 holds. Under the maximal positive correlation, there exist threshold values $\underline{S} = \underline{S}(E, V)$ and $\overline{S} = \overline{S}(E, V)$ with $0 < \underline{S} < \overline{S}$ such that a salient thinker selects a portfolio with*

$$\alpha \in \begin{cases} [0, \frac{1}{2}] & \text{if } S < \underline{S}, \\ \{\frac{1}{2}\} & \text{if } \underline{S} \leq S \leq \overline{S}, \\ [\frac{1}{2}, 1] & \text{if } S > \overline{S}. \end{cases}$$

The threshold values $\overline{S}$ and $\underline{S}$ are implicitly defined by the unique solutions to

$$\frac{2\sqrt{S^2+4} - 4S}{\sqrt{S^2+4} \cdot (\sqrt{S^2+4} - S)} = \pm \frac{\sqrt{V}}{E}$$

on the intervals $(\frac{2}{3}\sqrt{3}, \infty)$ and $(0, \frac{2}{3}\sqrt{3})$, respectively, and can be solved for numerically.

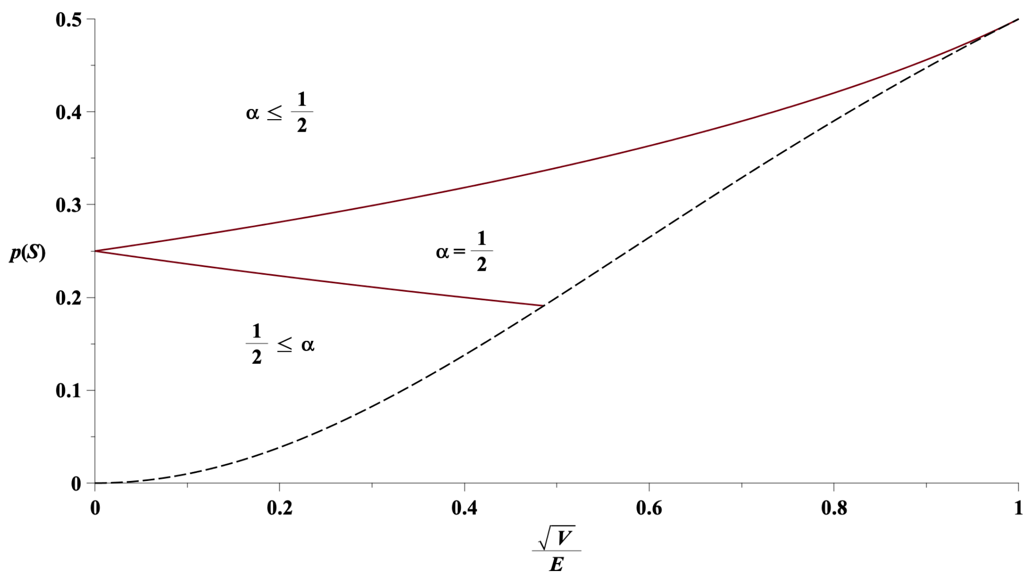

Figure 2: *Illustration of Proposition 3.*

3.3 Portfolio Selection with Symmetric Assets

The portfolio selection problems analyzed in this section are a generalization of the experimental design by Eyster and Weizsäcker (2016). In contrast to the previous section, we consider portfolios consisting of symmetric assets, and we start with the case of a continuous choice set (as in Eyster and Weizsäcker, 2016) before moving on to binary choice sets. The latter allows us to distinguish between salience theory and naive decision rules such as the $1/N$-heuristic.

Continuous choice sets. To guide our theoretical analysis, we propose a generalization of the paired portfolio selection problems (or twin problems) used in Eyster and Weizsäcker (2016) that captures the qualitative features of the numerical examples they used.[2]

Definition 6 (Generalized Twin Problem)**.** *Let* $\gamma \in [0, \frac{E}{\sqrt{V}}]$ *and consider the assets* X_1 *and* X_2 *with*

Asset / Prob.	$\frac{1}{2}$	$\frac{1}{2}$
X_1	$E + \sqrt{V}$	$E - \sqrt{V}$
X_2	$E - \gamma\sqrt{V}$	$E + \gamma\sqrt{V}$

A generalized twin problem *then consists of the following two investment decisions,*

$$(P.1) \quad \gamma = 1 \;\; and \;\; A = [0,1] \quad \& \quad (P.2) \quad \gamma \in \left[0, \frac{E}{\sqrt{V}}\right] \;\; and \;\; A = \left[\max\left\{0, \frac{\gamma - 1}{\gamma + 1}\right\}, 1\right],$$

where the choice set is given by $\mathcal{C} = \{X(\alpha) : \alpha \in A\}$ *in each of the problems.*

As the set of attainable portfolios is identical in both problems, any consequentialist theory of choice under risk (e.g., expected utility theory) predicts that, in a generalized twin problem, subjects choose the same portfolio in (P.1) and (P.2)—which requires the subject to adjust the composition of her portfolio in the second problem. The $1/N$-heuristic, in contrast, predicts naive diversification in both, (P.1) and (P.2), which implies different portfolios in the two problems. Also a salient thinker does not necessarily choose the same portfolio in both problems: while she always chooses the diversified portfolio in (P.1), this is not true in general for (P.2).

Proposition 4. *Consider a generalized twin problem. In* (*P.1*) *a salient thinker always chooses the diversified portfolio with* $\alpha = \frac{1}{2}$*, while in* (*P.2*) *she chooses a potentially under-diversified portfolio with*

$$\alpha \in \begin{cases} \{0, \frac{1}{2}\} & \text{if } \gamma = 0, \\ (0, \frac{\gamma}{1+\gamma}) & \text{if } 0 < \gamma < \frac{1}{3}, \\ (\frac{3\gamma - 1}{2\gamma + 2}, \frac{\gamma}{1+\gamma}) & \text{if } \frac{1}{3} \leq \gamma < 1, \\ \{\frac{\gamma}{1+\gamma}\} & \text{otherwise.} \end{cases}$$

[2] To be precise, Eyster and Weizsäcker (2016) implement joint distributions with four states of the world, and they have eight pairs of portfolio selection tasks where only six of these pairs are covered by our definition.

A salient thinker's behavior in a generalized twin problem is determined by diminishing sensitivity. Consider (P.1) and let $\alpha > \frac{1}{2}$. This portfolio's downside in state $(E - \sqrt{V}(2\alpha - 1), E)$ is more salient than its upside in state $(E + \sqrt{V}(2\alpha - 1), E)$ due to diminishing sensitivity. By symmetry, also for $\alpha < \frac{1}{2}$ the portfolio's downside is salient. Since the reference portfolio, which is given by $\alpha = \frac{1}{2}$, has neither a salient downside nor a salient upside it will be chosen. Next, consider (P.2) and, for the sake of the argument, suppose that $\gamma = 0$ holds, in which case the two states are $(E + \alpha\sqrt{V}, E + \frac{1}{2}\sqrt{V})$ and $(E - \alpha\sqrt{V}, E - \frac{1}{2}\sqrt{V})$. Again, whenever $\alpha \neq \frac{1}{2}$, the portfolio's downside is salient due to diminishing sensitivity. Hence, the salient thinker chooses either the reference portfolio with $\alpha = \frac{1}{2}$ or she chooses the diversified portfolio with $\alpha = 0$, thereby avoiding any variation in outcomes. By similar arguments, diminishing sensitivity implies that for any $\gamma \in (0, 1)$ there exist portfolios—different from the reference portfolio—with a salient upside, while for any $\gamma \geq 1$ only a portfolio's downside can be salient, which then implies that only the reference portfolio is attractive to a salient thinker.

Binary choice sets. To distinguish between salience theory and naive decision rules such as the $1/N$-heuristic, we consider a binary version of the generalized twin problems.

Definition 7 (Binary Twin Problem)**.** *Let* $\gamma \in [0, 3]$. *Consider the assets* X_1 *and* X_2 *with the joint distribution as in Definition 4. The* binary twin problem *then consists of the two investment decisions,*

$$(P.1) \quad \gamma = 1 \;\; \textit{and} \;\; A = \left\{ \frac{3 - \tilde{\gamma}}{4}, \frac{1}{2} \right\} \quad \& \quad (P.2) \quad \gamma = \tilde{\gamma} \neq 1 \;\; \textit{and} \;\; A = \left\{ \frac{1}{2}, \frac{\tilde{\gamma}}{1 + \tilde{\gamma}} \right\},$$

where the choice set is given by $\mathcal{C} = \{X(\alpha) : \alpha \in A\}$ *in each of the problems.*

In any binary twin problem, we adjust the set of admissable investments in X_1 in a way that (1) the set of resulting portfolios is exactly the same in (P.1) and (P.2), and (2) the agent can naively diversify in both problems. More specifally, we always offer the diversified portfolio given by $\alpha = \frac{\gamma}{1+\gamma}$ and the naively diversified portfolio given by $\alpha = \frac{1}{2}$.

Since the choice set is binary and identical in both (P.1) and (P.2), also the reference portfolio and a salient thinker's preferences over the available portfolios coincide. Hence, a salient thinker chooses the exact same portfolio in (P.1) and in (P.2). Diminishing sensitivity further implies that, in a binary twin problem, a salient thinker selects the diversified portfolio in both problems.

Proposition 5. *In any binary twin problem, a salient thinker diversifies in both,* (P.1) *and* (P.2)*.*

4 An Experiment on Portfolio Selection

4.1 Design and Sample Size

We plan to test for the salience predictions on portfolio selection with skewed and symmetric assets in an incentivized lab experiment. We use a between-subjects design with multiple decisions per subject, where in the treatment *Skewed* subjects make decisions that involve skewed assets (as illustrated in Table 2) and in the treatment *Symmetric* subjects make decisions that involve symmetric assets (as illustrated in Table 3).

Left-skewed Lottery	Right-skewed Lottery	E	V	$E/\sqrt{V}$	S	A
(120, 90%; 0, 10%)	(96, 90%; 216, 10%)	108	1296	3	2.7	$\{\alpha, 0.50\}$ or $[0, 1]$
(135, 64%; 60, 36%)	(81, 64%; 156, 36%)	108	1296	3	0.6	$\{\alpha, 0.50\}$ or $[0, 1]$
(40, 90%; 0, 10%)	(32, 90%; 72, 10%)	36	144	3	2.7	$\{\alpha, 0.50\}$ or $[0, 1]$
(45, 64%; 20, 36%)	(27, 64%; 52, 36%)	36	144	3	0.6	$\{\alpha, 0.50\}$ or $[0, 1]$
(80, 90%; 0, 10%)	(64, 90%; 144, 10%)	72	576	3	2.7	$\{\alpha, 0.50\}$ or $[0, 1]$
(90, 64%; 40, 36%)	(54, 64%; 104, 36%)	72	576	3	0.6	$\{\alpha, 0.50\}$ or $[0, 1]$

Table 2: *Mao pairs used in the treatment with skewed assets, whereby $\alpha \in \{0.10, 0.25, 0.75, 0.90\}$ is randomized at the subject-Mao-pair level. Each Mao pair is implemented under both, the perfectly negative and the maximal positive correlation.*

Each treatment consists of three (main) blocks: (1) subjects make 12 choices between exactly two portfolios, (2) subjects make 12 choices from a continuous set of portfolios, and (3) subjects make the exact same choices as in the first block, with the one exception that the portfolios are presented as reduced lotteries.[3] We randomize the order of blocks as well as the order of tasks within each block at the subject level. At the beginning of Block 1 and Block 2, subjects have to answer control questions to check for comprehension of the tasks. After the last decision of the last block (which could be either Block 1 or Block 2 or Block 3), subjects have to do a suprise memory task, in which they have to recall the joint distribution of the assets presented on the preceding screen. Subsequently, subjects have to answer financial literacy and modified CRT questions, and they have to solve math problems that are similar to computing the outcomes of a portfolio. Finally, we ask for demographics (incl. age, gender, field of study).

X_1	X_2 in (P.2)	E	V	$\tilde{\gamma}$	A in (P.1)	A in (P.2)
(96; 48)	(72; 72)	72	576	0	$\{0.75, 0.50\}$ or $[0, 1]$	$\{0.50, 0\}$ or $[0, 1]$
(79; 51)	(65; 65)	65	196	0	$\{0.75, 0.50\}$ or $[0, 1]$	$\{0.50, 0\}$ or $[0, 1]$
(120; 48)	(72; 96)	84	1296	0.33	$\{0.67, 0.50\}$ or $[0, 1]$	$\{0.50, 0.25\}$ or $[0, 1]$
(82; 58)	(66; 74)	70	144	0.33	$\{0.67, 0.50\}$ or $[0, 1]$	$\{0.50, 0.25\}$ or $[0, 1]$
(128; 72)	(16; 184)	100	784	3	$\{0.50, 0\}$ or $[0, 1]$	$\{0.75, 0.50\}$ or $[0.5, 1]$
(71; 49)	(27; 93)	60	121	3	$\{0.50, 0\}$ or $[0, 1]$	$\{0.75, 0.50\}$ or $[0.5, 1]$

Table 3: *Twin problems used in the treatment with symmetric assets, where (x; y) := (x, 50%; y, 50%).*

During the experiment the outcomes of the assets are presented in units of an experimental currency (ECU). At the end of the experiment, one of the 36 decisions in Skewed or 30 decisions in Symmetric, respectively, will be randomly drawn by the computer to be payoff-relevant. The outcome of the payoff-relevant choice will be also determined by a computer. In addition, subjects will be rewarded for their performance in the memory task (1 ECU per correctly remember outcome) as well as for correct answers to the control, financial literacy, modified CRT, and math

[3]Since the twin problems are constructed in a way that the portfolios are identical in (P.1) and (P.2), the third block includes only 6 decisions in the treatment with symmetric risks. Notice that, when computing the reduced lotteries that correspond to the portfolios presented in the Block 1, we round all outcomes on full numbers.

questions (1 ECU per correctly answered question). Experimental earnings will be converted at a ratio of 4 ECU : 1 Euro. Additionally the subjects receive a show-up fee of 4 Euro.

We plan to conduct 10 sessions with a total number of $n = 300$ subjects (150 subjects in the treatment Skewed and 150 subjects in the treatment Symmetric). The sessions will take place in January and February 2020 at the experimental laboratory of the University of Cologne.

4.2 Salience Predictions

In this section, we translate our general theoretical results into specific predictions given the parameters that we use in the experiment. In addition, we assume that a subject implements her intended choice—as predicted by our model—with some noise: with probability $\epsilon < \frac{1}{2}$, a subject trembles and does not take the option that maximizes her utility. Specifically, we assume that subjects maximize the salience-weighted utility subject to this implementation noise.

Portfolio selection with skewed assets. Suppose that the two assets correspond to one of the Mao pairs in Table 2. To begin with, we consider the binary choice between the diversified and some under-diversified portfolio. Proposition 1 yields our first prediction.

Prediction 1 (Binary, General)**.** *Let the two assets be perfectly negatively correlated.*

(*a*) *If the under-diversified portfolio is left-skewed, a majority of subjects chooses the diversified one.*

(*b*) *Weakly more subjects choose a right-skewed portfolio for $S = 2.7$ than for $S = 0.6$.*

If we impose Assumption 1, we obtain a stronger prediction, which will allow us to test this additional assumption. In sum, Assumption 1 implies that the selected portfolios vary with the correlation structure for $S = 0.6$, but not for $S = 2.7$. Since all Mao pairs in Table 2 have the same ratio $E/\sqrt{V}$, the salience predictions are identical across these Mao pairs.

Prediction 2 (Binary, Homogeneity of Degree Zero)**.** *Suppose that Assumption 1 holds.*

(*a*) *If $S = 0.6$, then, under the perfectly negative (maximal positive) correlation, a majority of subjects chooses the more (less) skewed portfolio, i.e., a majority of subjects chooses the under-diversified portfolio if and only if it is right-skewed (left-skewed).*

(*b*) *If $S = 2.7$, then a majority of subjects chooses the more skewed portfolio, i.e., a majority of subjects chooses the under-diversified portfolio if and only if it is right-skewed.*

(*c*) *The share of subjects choosing the more skewed portfolio is larger under the perfectly negative than under the maximal positive correlation if and only if $S = 0.6$.*

Under the perfectly negative correlation, our model also predicts the choice from a continuous set of portfolios consisting of the exact same Mao pairs as before. If the assets are maximally positively correlated, however, the salience model does not make a prediction.

Prediction 3 (Continuous, General)**.** *Let the assets be perfectly negatively correlated.*

(*a*) *A majority of subjects chooses either the diversified or a right-skewed, under-diversified portfolio.*

(*b*) *Weakly more subjects choose a right-skewed portfolio for $S = 2.7$ than for $S = 0.6$.*

Portfolio selection with symmetric assets. Suppose that the assets constitute a (binary) twin problem, as introduced in the previous section. First, we consider only binary twin problems. Proposition 5 together with our assumption on noise yields the following prediction.

Prediction 4 (Twin, Binary)**.**

(a) *In both, (P.1) and (P.2), a majority of subjects chooses the diversified portfolio.*

(b) *The share of subjects choosing the diversified portfolio is the same in (P.1) and in (P.2).*

Second, we consider only the generalized twin problems (with a continuous choice set). Proposition 4 together with our assumption on noise yields the following prediction.

Prediction 5 (Twin, Continuous)**.**

(a) *In (P.1), a majority of subjects chooses the diversified portfolio.*

(b) *In (P.2), a majority of subjects chooses a portfolio with $\alpha \in \{0, \frac{1}{2}\}$ if $\gamma = 0$, $\alpha \in (0, \frac{1}{4})$ if $\gamma = \frac{1}{3}$, or $\alpha = \frac{3}{4}$ if $\gamma = 3$, where α denotes the investment into the asset that is available also in (P.1).*

Additional predictions. Our salience model builds on the assumption that subjects perfectly integrate the two assets (i.e., they perfectly take the distribution of final payoffs into account). This results in the following prediction, which applies to both treatments

Prediction 6 (Binary, Reduced)**.** *The share of subjects choosing a given reduced lottery in the binary choice task (Block 3) is the same as the share of subjects choosing the portfolio with the exact same distribution over final payoffs in the binary portfolio selection task (Block 1).*

Finally, we want to understand how skewness shapes the perception and eventually the memory of outcomes. One could hypothesize that the extreme outcomes of less symmetric lotteries are more salient and therefore more precisely remembered. This could be interpreted as subjects being the more likely to think about the final outcomes of a portfolio the less symmetric the underlying lotteries are. Since these considerations are outside of our salience model, we view this part of the experiment as exploratory.

4.3 Statistical Analysis

In this section, we describe the statistical methods that we will use to test for our predictions. Since each subject makes multiple decisions in the experiment, we will estimate simple OLS regressions, which allow us to properly cluster the standard errors at the subject level. Let I_{Skew} and I_{Sym} be the sets of subjects in the treatments Skewed and Symmetric, respectively.

Test of Prediction 1: We construct a binary indicator $D_{i,k}$ that takes a value of one if a subject $i \in I_{Skew}$ has chosen the diversified portfolio in decision $k \in \{1, \ldots, 12\}$ of Block 1 and a value of zero otherwise. We use only the subsample of choices in which a subject faces a negatively correlated Mao pair. To test for Part (a), we consider for each subject the subsample of choices that include a left-skewed portfolio. On this subsample, we regress $D_{i,k} - \frac{1}{2}$ on a constant. If the constant is not significantly larger than zero, we reject our hypothesis. To test for Part (b), we construct a binary indicator $Skewed_{i,k}$ that takes a value of one if subject $i \in I_{Skew}$ faces in decision $k \in \{1, \ldots, 12\}$ a Mao pair with $S = 2.7$ and a value of zero otherwise. Then, we regress $D_{i,k}$ on $Skewed_{i,k}$. If the coefficient on Skewed is positive and statistically significant, we reject our hypothesis.

Test of Prediction 2: We construct a binary indicator $M_{i,k}$ that takes a value of one if a subject $i \in I_{Skew}$ has chosen the more skewed portfolio in decision $k \in \{1, \ldots, 12\}$ of Block 1 and a value of zero otherwise. To test for Part (a), we use only the subsample of choices with $Skewed_{i,k} = 0$. Here, we regress $M_{i,k} - \frac{1}{2}$ on a constant, separately for the subsamples of choices with the perfectly negative and the maximal positive correlation. We reject our hypothesis if either the constant for the perfectly negative (maximal positive) correlation is not significantly larger (smaller) than zero. To test for Part (b), we use only the subsample of choices with $Skewed_{i,k} = 1$, and we regress $M_{i,k} - \frac{1}{2}$ on a constant, separately for the subsamples of choices with the perfectly negative and the maximal positive correlation. We reject our hypothesis if not both constants are significantly larger than zero. Finally, to test for Part (c), we introduce a binary indicator $Negative_{i,k}$ which takes a value of one if subject $i \in I_{Skew}$ faces a negatively correlated Mao pair in decision $k \in \{1, \ldots, 12\}$ and a value of zero otherwise. Then, we regress $M_{i,k}$ on $Negative_{i,k}$, separately for the subsample with $Skewed_{i,k} = 1$ and the subsample with $Skewed_{i,k} = 0$. We reject our hypothesis if we estimate a statistically significant coefficient for the subsample with $Skewed_{i,k} = 1$ or if the coefficient is either insignificantly different from zero or negative for the subsample with $Skewed_{i,k} = 0$.

Test of Prediction 3: We construct a binary indicator $WR_{i,k}$ that takes a value of one if a subject $i \in I_{Skew}$ has chosen either the diversified or a right-skewed portfolio in decision $k \in \{1, \ldots, 12\}$ of Block 2 and a value of zero otherwise. We use only the subsample of choices in which a subject faces a negatively correlated Mao pair. To test for Part (a), we regress $WR_{i,k} - \frac{1}{2}$ on a constant. We reject our hypothesis if the constant is either negative or positive but not statistically significant. To test for Part (b), we introduce the binary indicator $R_{i,k}$ that takes a value of one if a subject $i \in I_{Skew}$ has chosen a right-skewed portfolio in decision $k \in \{1, \ldots, 12\}$ of Block 2 and a value of zero otherwise. Then, we regress $R_{i,k}$ on $Skewed_{i,k}$. If the coefficient on Skewed is negative and significant, we reject our hypothesis.

Test of Prediction 4: We construct a binary indicator $DT_{i,k}$ that takes a value of one if a subject $i \in I_{Sym}$ has chosen the diversified portfolio in decision $k \in \{1, \ldots, 12\}$ of Block 1 and a value of zero otherwise. In addition, we construct a binary indicator $P1_{i,k}$ that takes a value of one if a subject $i \in I_{Sym}$ faces the problem (P.1) of a given twin problem in decision $k \in \{1, \ldots, 12\}$ of Block 1 and a value of zero otherwise. To test for Part (a), we regress $DT_{i,k} - \frac{1}{2}$ on a constant, separately for the subsample with $P1_{i,k} = 1$ and the subsample with $P1_{i,k} = 0$. We reject our hypothesis if one of these constants is not significantly positive. To test for Part (b), we regress $DT_{i,k}$ on $P1_{i,k}$. If the coefficient on $P1_{i,k}$ is significant, we reject our hypothesis.

Test of Prediction 5: We construct a binary indicator $DTC_{i,k}$ that takes a value of one if a subject $i \in I_{Sym}$ has chosen the diversified portfolio in decision $k \in \{1, \ldots, 12\}$ of Block 2 and a value of zero otherwise. To test for Part (a), we regress $DTC_{i,k} - \frac{1}{2}$ on a constant. We reject our hypothesis unless the constant is positive and statistically significant. To test for Part (b), we contruct a binary indicator $SC_{i,k}$ that takes a value of one if a subject $i \in I_{Sym}$ has chosen a portfolio consistent with salience theory in decision $k \in \{1,\ldots,12\}$ of Block 2 and a value of zero otherwise. We then regress $SC_{i,k} - \frac{1}{2}$ on a constant, and we reject our hypothesis if the constant not significantly larger than zero.

Test of Prediction 6: To save on notation, we describe the test only for the treatment with skewed assets. Here, we construct a binary indicator $DR_{i,k}$ that takes a value of one if a subject $i \in I_{Skew}$ has chosen the reduced lottery with a lower variance in decision $k \in \{1,\ldots,12\}$ of Block 3 and a value of zero otherwise. Let $\widetilde{DR}_{i,k}$ be a permutation of the vector $DR_{i,k}$ sorted as follows: The first and the second entry correspond to the Mao pair with the lowest expected value and $S = 0.6$ under the perfectly negative correlation and the maximal positive correlation, respectively. The third and the fourth entry correspond to the Mao pair with the lowest expected value and $S = 2.7$ under the perfectly negative correlation and the maximal positive correlation, respectively. The fifth and the sixth entry correspond to the Mao pair with the second-lowest expected value and $S = 0.6$ under the perfectly negative correlation and the maximal positive correlation, respectively, etc. In addition, let $\widetilde{D}_{i,k}$ be a permutation of the vector $D_{i,k}$ sorted in the same way. Then, we regress $\widetilde{DR}_{i,k} - \widetilde{D}_{i,k}$ on a constant, and we reject our hypothesis if the constant is statistically significant.

References

Benartzi, S. and Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American economic review*, **91** (1), 79–98.

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

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

—, Nicola Gennaioli and Shleifer, A. (2016). Competition for attention. *Review of Economic Studies*, **83** (2), 481–513.

Dertwinkel-Kalt, M. and Köster, M. (forthcoming). Salience and skewness preferences. *Journal of the European Economic Association*.

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

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

Eyster, E. and Weizsäcker, G. (2016). Correlation neglect in portfolio choice: Lab evidence. *Working Paper*.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47** (2), 263–291.

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

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, **7** (1), 77–91.

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

Appendix A: Proofs

A.1: Preliminary Results

Lemma 3. *In the generalized twin problem,* (P.1) *and* (P.2) *implement the exact same set of portfolios, and the diversified portfolio satisfies* $\alpha = \frac{1}{2}$ *in* (P.1) *and* $\alpha = \frac{\gamma}{1+\gamma}$ *in* (P.2).

Proof. Denote the investment in X_1 as α_1 in (P.1) and as α_2 in (P.2), respectively.

In a first step, we show that *(P.1) and (P.2) indeed implement the same set of portfolios.* Notice that in (P.1), for any $x \in [0, \frac{1}{2}]$, the investments $\alpha_1 = \frac{1}{2} - x$ and $\alpha_1 = \frac{1}{2} + x$ implement the same portfolio, and that, for any $\alpha_1 \in [\frac{1}{2}, 1]$, we can implement the same portfolio in (P.2) by setting

$$\alpha_2 = \frac{\gamma + 2\alpha_1 - 1}{\gamma + 1}.$$

This further implies that we have to restrict the choice set in (P.2) to be

$$A = \left[\max\left\{\frac{\gamma - 1}{\gamma + 1}, 0\right\}, 1\right],$$

as otherwise (P.2) would offer larger set of attainable portfolios than (P.1).

In a second step, we show that *the diversified portfolio satisfies* $\alpha = \frac{\gamma}{1+\gamma}$. Since X_1 and X_2 have the same expected value, we simply have to find the portfolio with the minimal variance:

$$Var(\alpha X_1 + (1-\alpha)X_2) = \alpha^2 Var(X_1) + (1-\alpha)^2 Var(X_2) + 2\alpha(1-\alpha)Cov(X_1, X_2).$$

Since $Var(X_1) = V$, $Var(X_2) = \frac{1}{2}V(1+\gamma^2)$, and $Cov(X_1, X_2) = \frac{1}{2}V(1-\gamma)$, we have

$$Var(\alpha X_1 + (1-\alpha)X_2) = V\left(\alpha^2 + (1-\alpha)^2\frac{1}{2}(1+\gamma^2) + \alpha(1-\alpha)(1-\gamma)\right),$$

and therefore

$$\begin{aligned} \frac{\partial}{\partial\alpha} Var(\alpha X_1 + (1-\alpha)X_2) &= 2\alpha - (1-\alpha)(1+\gamma^2) + (1-\alpha)(1-\gamma) - \alpha(1-\gamma) \\ &= -(1-\alpha)\gamma(1+\gamma) + \alpha(1+\gamma) \\ &= \alpha(1+\gamma)^2 - \gamma(1+\gamma) \end{aligned} \tag{3}$$

which is equal to zero if and only if $\alpha = \frac{\gamma}{1+\gamma}$. The claim then follows from the fact that, by (3), $Var(\alpha X_1 + (1-\alpha)X_2)$ is obviously convex in α. □

A.2: Portfolio Selection with Skewed Assets

Proof of Lemma 2. Let $X_1 = L(E, V, S)$ and $X_2 = L(E, V, -S)$. Then, we have

$$\begin{aligned} Var(\alpha X_1 + (1-\alpha)X_2) &= \alpha^2 V + (1-\alpha)^2 V + 2\alpha(1-\alpha)Cov(X_1, X_2) \\ &= V - 2\alpha(1-\alpha)\big(V - Cov(X_1, X_2)\big) \\ &= V\big[1 - 2\alpha(1-\alpha)\big(1 - \rho(X_1, X_2)\big)\big], \end{aligned}$$

where the second equality follows from the definition of the correlation coefficient and fact that X_1 and X_2 have the same variance. Since the correlation of lotteries of a Mao pair cannot exceed a certain threshold $\overline{\rho} < 1$, it follows that the variance of the portfolio is minimized at $\alpha = \frac{1}{2}$. Moreover, for a fixed correlation, the variance is symmetric around $\alpha = \frac{1}{2}$. The additional statements regarding the perfectly negative correlation are straightforward to prove. □

Proof of Proposition 1. Lemma 2 & Proposition 3 in Dertwinkel-Kalt and Köster (forthcoming). □

Proof of Corollary 1. PART (a): By Proposition 1, we can restrict our attention to portfolios with $\alpha > \frac{1}{2}$. Under the perfectly negative correlation, the portfolio $X(\alpha)$ thus implements the binary lottery $L(E, (2\alpha-1)^2 V, S)$, while the alternative $X(\frac{1}{2})$ pays E with certainty. A salient thinker thus chooses $X(\alpha)$ over $X(\frac{1}{2})$ if and only if the upside of $L(E, (2\alpha-1)^2 V, S)$ is salient, which—under Assumption 1—is indeed the case if and only if

$$\frac{x_2(E, (2\alpha-1)^2 V, S)}{E} > \frac{E}{x_1(E, (2\alpha-1)^2 V, S)},$$

or, equivalently,

$$\alpha < \frac{1}{2} + \frac{ES}{\sqrt{4V}}.$$

PART (b): Let $\alpha < \frac{1}{2}$. Suppose that the two assets are maximally positively correlated, in which case the salient thinker chooses between the following two portfolios:

	p	$1-2p$	p
$X(\alpha)$	$E - \frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}$	$E - \frac{\sqrt{V}p[2\alpha-1]}{\sqrt{p(1-p)}}$	$E + \frac{\sqrt{V}[\alpha(1-2p)+p]}{\sqrt{p(1-p)}}$
$X(\frac{1}{2})$	$E - \frac{\sqrt{V}}{2\sqrt{p(1-p)}}$	E	$E + \frac{\sqrt{V}}{2\sqrt{p(1-p)}}$

It is easy to check that a salient thinker prefers $X(\alpha)$ over $X(\frac{1}{2})$ if and only if

$$(1-2\alpha) \cdot K(\alpha, p) > 0.$$

where the second term is defined as follows

$$K(\alpha,p) := \sigma\bigg(E-\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}, E-\frac{\sqrt{V}}{2\sqrt{p(1-p)}}\bigg)$$
$$+\sigma\bigg(E+\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}, E+\frac{\sqrt{V}}{2\sqrt{p(1-p)}}\bigg) - 2\sigma\bigg(E-\frac{\sqrt{V}p[2\alpha-1]}{\sqrt{p(1-p)}}, E\bigg).$$

By diminishing sensitivity, we have

$$\sigma\bigg(E-\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}, E-\frac{\sqrt{V}}{2\sqrt{p(1-p)}}\bigg)$$
$$> \sigma\bigg(E+\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}, E+\frac{\sqrt{V}}{2\sqrt{p(1-p)}}\bigg),$$

so that a sufficient condition for $K(\alpha,p) > 0$ is given by

$$\sigma\bigg(E+\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}, E+\frac{\sqrt{V}}{2\sqrt{p(1-p)}}\bigg) \geq \sigma\bigg(E-\frac{\sqrt{V}p[2\alpha-1]}{\sqrt{p(1-p)}}, E\bigg).$$

By Assumption 1, the preceding inequality holds for $\alpha < \frac{1}{2}$ if and only if

$$\frac{E+\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}}{E+\frac{\sqrt{V}}{2\sqrt{p(1-p)}}} \geq \frac{E-\frac{\sqrt{V}p[2\alpha-1]}{\sqrt{p(1-p)}}}{E},$$

or, equivalently,

$$\alpha \geq \frac{E}{\sqrt{V}}\sqrt{\frac{1-p}{p}}\left(\frac{1-4p}{2-4p}\right) + \frac{1-p}{1-2p}.$$

By the same arguments, a sufficient condition for $K(\alpha,p) < 0$—and thus for a salient thinker to choose $X(\frac{1}{2})$ over $X(\alpha)$ for $\alpha < \frac{1}{2}$—, is that

$$\sigma\bigg(E-\frac{\sqrt{V}p[2\alpha-1]}{\sqrt{p(1-p)}}, E\bigg) \geq \sigma\bigg(E-\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}, E-\frac{\sqrt{V}}{2\sqrt{p(1-p)}}\bigg).$$

By Assumption 1, the preceding inequality holds for $\alpha < \frac{1}{2}$ if and only if

$$\frac{E-\frac{\sqrt{V}p[2\alpha-1]}{\sqrt{p(1-p)}}}{E} \geq \frac{E-\frac{\sqrt{V}}{2\sqrt{p(1-p)}}}{E-\frac{\sqrt{V}[\alpha(2p-1)+(1-p)]}{\sqrt{p(1-p)}}},$$

or, equivalently,

$$\alpha \leq \frac{E}{\sqrt{V}}\sqrt{\frac{1-p}{p}}\left(\frac{1-4p}{2-4p}\right) - \frac{p}{1-2p}.$$

The arguments for $\alpha > \frac{1}{2}$ are analoguous and therefore omitted. □

Proof of Proposition 2. Lemma 2 & Proposition 3 in Dertwinkel-Kalt and Köster (forthcoming). □

Proof of Corollary 2. The proof goes along the same lines as that of Corollary 1 (a). □

Proof of Proposition 3. The proof goes pretty much along the lines as that of Corollary 1 (b). We therefore only sketch the differences here. For any portfolio $\alpha > \frac{1}{2}$, there is only one upside relative to the diversified portfolio, which is, by construction, valued at its expected value. If this upside is least salient, then $X(\alpha)$ is valued below its expected value and is therefore less attractive than the diversified portfolio. Under Assumption 1, the upside is indeed least salient if and only if $S \leq \overline{S}$. Analogously, for $\alpha < \frac{1}{2}$, there is only one downside relative to the diversified portfolio, but if this downside is most salient, then $X(\alpha)$ is valued less than its expected value. Under Assumption 1, the downside is indeed most salient if and only if $S \geq \underline{S}$. □

A.3: Portfolio Selection with Symmetric Assets

Proof of Proposition 4. It is sufficient to find portfolios that yield a salience-weighted utility that (weakly) exceeds the expected value of E. Recall that, by construction, the reference portfolio yields a salience-weighted utility of E. We subsequently analyze (P.1) and (P.2).

1. STEP: Consider (P.1). Here, we have $U^s(X(\alpha)|\mathcal{C}) \geq E$ if and only if

$$(2\alpha - 1)\big[\sigma(E + \sqrt{V}(2\alpha - 1), E) - \sigma(E - \sqrt{V}(2\alpha - 1), E)\big] \geq 0.$$

By diminishing sensitivity, we have

$$\sigma(E + \sqrt{V}(2\alpha - 1), E) > \sigma(E - \sqrt{V}(2\alpha - 1), E) \quad \iff \quad 2\alpha - 1 < 0,$$

which in turn implies that $U^s(X(\alpha)|\mathcal{C}) \geq E$ if and only if $\alpha = \frac{1}{2}$.

2. STEP: Consider (P.2). If $\gamma < 1$, then $U^s(X(\alpha)|\mathcal{C}) > E$ if and only if

$$(\alpha - (1-\alpha)\gamma)\left[\sigma(E + \sqrt{V}(\alpha - (1-\alpha)\gamma), E + \frac{1-\gamma}{2}\sqrt{V}) - \sigma(E - \sqrt{V}(\alpha - (1-\alpha)\gamma), E - \frac{1-\gamma}{2}\sqrt{V})\right] > 0,$$

or, equivalently,

$$\alpha \in \left(\frac{3\gamma - 1}{2\gamma + 1}, \frac{\gamma}{1+\gamma}\right).$$

If $\gamma \geq 1$, then $U^s(X(\alpha)|\mathcal{C}) > E$ if and only if

$$\big[\alpha - (1-\alpha)\gamma\big]\left[\sigma(E + \sqrt{V}(\alpha - (1-\alpha)\gamma), E) - \sigma(E - \sqrt{V}(\alpha - (1-\alpha)\gamma), E)\right] > 0,$$

which cannot hold due to diminishing sensitivity. This implies that $U^s(X(\alpha)|\mathcal{C})$ is maximized for $\alpha = \frac{\gamma}{1+\gamma}$, which was to be proven. □

Proof of Proposition 5. Since the choice set is binary and identical in both problems, the decision is exactly the same from a salient thinker's perspective and she therefore chooses the same portfolio in (P.1) and in (P.2). In other words, since the reference portfolio does not change across the two problems, also a salient thinkers preferences over the two portfolios do not change. Hence, it is sufficient to show that a salient thinker chooses the diversified portfolio in (P.1).

In (P.1), for $\tilde{\gamma} \in [0, 3]$, the salient thinker chooses between the following two portfolios:

Asset / Prob.	$\frac{1}{2}$	$\frac{1}{2}$
$X(\frac{1}{2})$	E	E
$X(\frac{3-\tilde{\gamma}}{4})$	$E + \frac{1}{2}\sqrt{V}(1-\tilde{\gamma})$	$E - \frac{1}{2}\sqrt{V}(1-\tilde{\gamma})$

The salient thinker thus prefers $X(\frac{1}{2})$ over $X(\frac{3-\tilde{\gamma}}{4})$ if and only if

$$\frac{1}{2}\sqrt{V}(1-\tilde{\gamma})\left[\sigma\left(E, E - \frac{1}{2}\sqrt{V}(1-\tilde{\gamma})\right) - \sigma\left(E, E + \frac{1}{2}\sqrt{V}(1-\tilde{\gamma})\right)\right] > 0,$$

which is indeed the case due to diminishing sensitivity. □