Tail Distortion Risk and Its Asymptotic Analysis

Li Zhu∗ Haijun Li†

May 2011
Revision: March 2012

Abstract

A distortion risk measure used in finance and insurance is defined as the expected value of potential loss under a scenario probability measure. In this paper, the tail distortion risk measure is introduced to assess tail risks of excess losses modeled by the right tails of loss distributions. The asymptotic linear relation between tail distortion and Value-at-Risk is derived for heavy tailed losses with the linear proportionality constant depending only on the distortion function and tail index. Various examples involving tail distortions for location, scale and shape invariant loss distribution families are also presented to illustrate the results.

JEL code and keywords: G32, distortion risk measure, regular variation, tail risk, tail conditional expectation.

Classification codes: IM10, IM54.

1 Introduction

Let $L$ be the convex cone\footnote{A subset $\mathcal{L}$ of a linear space is a convex cone if $x_1 \in \mathcal{L}$ and $x_2 \in \mathcal{L}$ imply that $\lambda_1 x_1 + \lambda_2 x_2 \in \mathcal{L}$ for any $\lambda_1 > 0$ and $\lambda_2 > 0.$} consisting of all the performance variables which represent loss of a financial portfolio at the end of a given period. Note that $-X$, where $X \in L$, represents the net worth of a financial position. To analyze right tail risks of loss distributions, one often uses the tail conditional expectation (TCE):

$$TCE_p(X) := E(X \mid X > \text{VaR}_p(X)),$$

(1.1)

∗lzhu@math.wsu.edu, Department of Mathematics, Washington State University, Pullman, WA 99164, U.S.A.
†lih@math.wsu.edu, Department of Mathematics, Washington State University, Pullman, WA 99164, U.S.A. Research is supported by NSF grants CMMI 0825960 and DMS 1007556.
where \( \text{VaR}_p(X) := \sup\{x \in \mathbb{R} : P(X > x) > 1 - p\} \) is known as the Value-at-Risk (VaR) with confidence level \( p \). VaR is the most commonly used risk measure in finance and insurance, but TCE focuses more than VaR does on assessing right tail risk. Since we are mainly interested in analyzing extremal risk for right tails of loss distributions, we assume throughout this paper that loss variables \( X \) are nonnegative.

Observe from (1.1) that in the case of continuous loss distributions, VaR and TCE can be easily expressed as follows.

\[
\text{VaR}_p(X) = \int_0^\infty g_{\text{VaR}}(P(X > x))dx, \\
\text{TCE}_p(X) = \int_0^\infty g_{\text{TCE}}(P(X > x))dx
\]

where \( g_{\text{VaR}}(x) := I\{x > 1 - p\} \) is the indicator function of \( \{x > 1 - p\} \), and \( g_{\text{TCE}}(x) := \min\{x/(1 - p), 1\} \). That is, both VaR and TCE are the risk measures that can be expressed in terms of Choquet integrals [3].

**Definition 1.1.** For a given nondecreasing function \( g : [0, 1] \mapsto [0, 1] \) such that \( g(0) = 0 \) and \( g(1) = 1 \), the Choquet integral \( H_g(X) \) of any nonnegative random variable \( X \) with distribution \( F_X \) is defined as follows:

\[
H_g(X) = \int_0^\infty g(1 - F_X(x))dx = \int_0^\infty g(F_X(x))dx,
\]

where \( F_X(x) = 1 - F_X(x) \) denotes the survival function of \( X \).

The function \( g \) in (1.2) is called a distortion function, and if \( X \) is viewed as a loss variable, then \( H_g(X) \) is also called a distortion risk measure of \( X \) with distortion function \( g \). Let \( \mathbb{P}_{g,F} \) denote the probability measure induced by the distribution function \( \bar{g}(F_X(x)) \), where \( \bar{g}(u) = 1 - g(1 - u) \), then

\[
H_g(X) = \int_0^\infty (1 - \bar{g}(F_X(x)))dx = \mathbb{E}_{\mathbb{P}_{g,F}}(X).
\]

That is, a distortion risk \( H_g(X) \) measures the expected value of \( X \) under the distorted probability measure \( \mathbb{P}_{g,F} \) that may describe a possible scenario. The basic properties of Choquet integrals (see [3], chapters 5 and 6) are listed below.

**Lemma 1.2.** Let \( g : [0, 1] \mapsto [0, 1] \) be a distortion function, and then the distortion risk measure \( H_g(\cdot) \) satisfies following properties.

1. (monotonicity) For \( X, Y \in \mathcal{L} \) with \( X \leq Y \) almost surely, \( H_g(X) \leq H_g(Y) \).
2. (positive homogeneity) For all $X \in \mathcal{L}$ and every $\lambda > 0$, $H_g(\lambda X) = \lambda H_g(X)$.

3. (translation invariance) For all $X \in \mathcal{L}$ and every $l \in \mathbb{R}$, $H_g(X + l) = H_g(X) + l$.

4. (subadditivity) If $g$ is concave, then $H_g(X + Y) \leq H_g(X) + H_g(Y)$.

5. (superadditivity) If $g$ is convex, then $H_g(X + Y) \geq H_g(X) + H_g(Y)$.

The monotonicity says a portfolio with more potential loss is riskier. The subadditivity is the property of risk reduction by diversification. The homogeneity describes how portfolio size directly influences its risk, whereas the translation invariance describes how incurring additional (deterministic) loss would affect portfolio risk.

Remark 1.3. 1. In general, a risk measure $\varrho$ is defined as a measurable mapping, with some basic operational axioms, from the space of all loss variables into $\mathbb{R}$, and these operational axioms reflect the risk perception of agents (or regulators) involved in the situation under consideration [8]. The risk $\varrho(X)$ for loss $X$ corresponds to the amount of extra capital requirement that has to be invested in some secure instrument so that the resulting position $\varrho(X) - X$ is acceptable to agents.

2. A risk measure related to distortion risk is the coherent risk measure introduced in [1]. A mapping $\varrho : \mathcal{L} \mapsto \mathbb{R}$ is called a coherent risk measure if $\varrho$ satisfies the economically coherent axioms of (1), (2), (3), and (4) in Lemma 1.2. Under some regularity conditions (such as the Fatou property, see [4]), a coherent risk measure $\varrho(X)$ arises as the supremum of expected values of loss $X$ under various scenarios:

$$\varrho(X) = \sup_{Q \in \mathcal{S}} \mathbb{E}_Q(X)$$

(1.3)

where $\mathbb{E}_Q(\cdot)$ denotes the expectation with respect to the probability measure $Q$, and $\mathcal{S}$ is a convex set of scenario probability measures on states, that are absolutely continuous with respect to the underlying measure $\mathbb{P}$. If the scenario set $\mathcal{S} = \{\mathbb{P}(\cdot | A) : \mathbb{P}(A) \geq 1 - p\}$, then $\varrho(X)$ is known as the worst conditional expectation, and in the case of continuous losses, $\varrho(X)$ equals to $TCE_p(X)$.

It follows from Lemma 1.2 that $H_g(X)$ is a coherent risk measure for any concave distortion function $g$, but there are also non-coherent distortion risk measures. For example, both VaR and TCE can be written in terms of Choquet integrals, hence, they are two distortion risk measures (see, e.g., [22, 23] for details). In contrast to TCE that is a coherent and distortion risk measure, VaR violates the subadditivity and thus is not coherent. Note that there are also non-distortion coherent risk measures (see [4]).
Early work on distortion risk measures and their representation can be found in [12, 24] and the references therein. The connection of distortion risk to coherent risk was first established in [16, 17], and early work on maxmin ordering preferences and their dual integral representation can be found in [9, 6]. Goovaerts et al. in a recent paper [8] explained the origin of distortion risk measures, and their use in a risk management setting. In particular, [8] highlights the role of axiomatic characterizations in specifying risk in the situation under consideration at the modeling level, and in deriving at the dual level a decision principle based on optimization procedures to quantify the risk. Distortion risk measures were introduced in the actuarial literature in [3, 19, 20]. The relations among risk, entropy and distortion were discussed in [13].

The goal of this paper is to study the asymptotic behavior of distortion risk measures focusing on the right tail of a heavy-tailed loss distribution. A non-negative loss variable \(X\) with distribution function \(F\) has a heavy or regularly varying right tail at \(\infty\) with tail index \(\alpha > 0\) if its survival function is of the following form (see, e.g., [2] for detail),

\[
\overline{F}(x) := \mathbb{P}(X > x) = x^{-\alpha}L(x), \quad x > 0, \quad \alpha > 0, \tag{1.4}
\]

where \(L\) is a slowly varying function; that is, \(L\) is a positive function on \((0, \infty)\) with property \(\lim_{x \to \infty} L(cx)/L(x) = 1\), for every \(c > 0\). The distributions (1.4) constitute a large semi-parametric family of loss distributions, including the Pareto and Fréchet distributions, that fit various data in finance and insurance [14]. While the distortion measures for loss distributions (1.4) generally do not have closed-form expressions, their asymptotic behaviors, as we will show in this paper, can be described explicitly in terms of the distortion function \(g\) and tail index \(\alpha\).

To study the asymptotic behavior of distortion risk, we introduce tail distortion risk measures.

**Definition 1.4.** For a given nondecreasing function \(g : [0, 1] \to [0, 1]\) such that \(g(0) = 0\) and \(g(1) = 1\), the tail distortion risk measure \(H_g(\cdot)\) of any nonnegative random variable \(X\) is defined as follows:

\[
H_g(X| X > \text{VaR}_p(X)) = \int_0^\infty g(\overline{F}_{X|X>\text{VaR}_p(X)}(x))dx, \tag{1.5}
\]

where \(\overline{F}_{X|X>t}(x) = 1 - F_{X|X>t}(x) = 1 - \mathbb{P}(X \leq x|X > t)\).

**Proposition 1.5.** The tail distortion risk of a continuous loss \(X\), as defined above, is a distortion risk measure.
Proof. Let $t = \text{VaR}_p(X)$, $0 < p < 1$, and consider

$$H_g(X | X > t) = \int_0^\infty g(P(X > x | X > t)) dx$$

$$= \int_0^\infty g \left( \frac{P(X > x, X > t)}{P(X > t)} \right) dx$$

$$= \int_0^t g(1) dx + \int_t^\infty g \left( \frac{P(X > x)}{P(X > t)} \right) dx$$

$$= \int_0^t dx + \int_t^\infty g \left( \frac{P(X > x)}{1 - p} \right) dx \quad (1.6)$$

Let’s define a function as follows:

$$g_p(u) = \begin{cases} 
  g\left( \frac{u}{1-p} \right) & \text{if } 0 \leq u < 1 - p \\
  1 & \text{if } 1 - p \leq u \leq 1 
\end{cases} \quad (1.7)$$

and then it is easy to see that $g_p$ is a nondecreasing function satisfying $g_p(0) = 0$, $g_p(1) = 1$. Combining (1.6) and (1.7), we obtain

$$H_g(X | X > \text{VaR}_p(X)) = \int_0^\infty g_p(P(X > x)) dx, \quad (1.8)$$

which is a distortion risk measure with distortion function $g_p$, $0 < p < 1$. \hfill \square

It follows from (1.8) that the tail distortion $H_g(X | X > \text{VaR}_p(X))$ of $X$ can be viewed as the expected value of $X$ under a change of the underlying measure that is deformed on the tail loss distribution. If $g = 1$, then $H_g(X | X > \text{VaR}_p(X)) = \text{TCE}_p(X)$. It is known (see, e.g., [10]) that for a heavy-tailed loss $X$ with distribution (1.4), the asymptotic relation of $\text{TCE}_p(X)$ and $\text{VaR}_p(X)$ can be derived as follows: for $\alpha > 1$,

$$\text{TCE}_p(X) \approx \frac{\alpha}{\alpha - 1} \text{VaR}_p(X), \quad p \to 1. \quad (1.9)$$

The main result of this paper, to be detailed in Section 2, establishes an explicit, asymptotic relation between $H_g(X | X > \text{VaR}_p(X))$ and $\text{VaR}_p(X)$, as $p \to 1$, which includes (1.9) as a special case. Our method is based on a limiting result for integrals of distortion functions with respect to regularly varying loss distributions. The tail distortions for location, scale, and shape invariant families are discussed in Section 3 to illustrate our asymptotic result. Finally, some comments and an example involving insurance premium pricing are discussed in Section 4 to conclude the paper.
2 Asymptotic Properties of Tail Distortion Risk

Let $X \geq 0$ denote a loss with regularly varying distribution (1.4). We use $RV_{-\alpha}$ in this paper to denote the class of all nonnegative loss variables whose survival functions are regularly varying with tail index $\alpha > 0$. Roughly speaking, regularly varying functions are those functions which behave asymptotically like power functions. The theory of regularly varying functions is an essential analytical tool for dealing with heavy tails, long-range dependence and domains of attraction, and the detailed discussions on these functions can be found in [2, 14]. In particular, the following uniform convergence for regularly varying functions can be found in [14], Section 2.1.

**Lemma 2.1.** For any regularly varying function $U(x) = x^{-\alpha}L(x)$ for $\alpha \in \mathbb{R}_+$, where $L(x)$ is slowly varying, one has that $\lim_{t \to \infty} U(tx)/U(t) = x^{-\alpha}$ uniformly on intervals of the form $(b, \infty)$, $b > 0$.

Observe from (1.6) that

$$H_g(X \mid X > t) = t + \int_t^\infty g \left( \frac{P(X > x)}{P(X > t)} \right) dx,$$

where $t = \text{VaR}_p(X)$. Let $x = wt$, we obtain that

$$H_g(X \mid X > t) = t + t \int_1^\infty g \left( \frac{P(X > wt)}{P(X > t)} \right) dw,$$

that is, for any $t > 0$,

$$\frac{H_g(X \mid X > t)}{t} = 1 + \int_1^\infty g \left( \frac{P(X > wt)}{P(X > t)} \right) dw. \quad (2.1)$$

Since $t = \text{VaR}_p(X) \to \infty$, as $p \to 1$, the asymptotic relation between $H_g(X \mid X > \text{VaR}_p(X))$ and $\text{VaR}_p(X)$, as $p \to 1$, boils down to the convergence of the integral in (2.1). We first illustrate our idea using continuous distortion functions.

**Proposition 2.2.** If distortion function $g(\cdot)$ is continuous, and a loss random variable $X$ is regularly varying (i.e., $X \in RV_{-\alpha}$, $\alpha > 0$), then

$$g \left( \frac{P(X > wt)}{P(X > t)} \right) \to g(w^{-\alpha}), \text{ as } t \to \infty,$$

uniformly on $[1, \infty)$. 

6
Proof. Since function $g(\cdot)$ is continuous on $[0, 1]$, then by the Heine-Cantor theorem, $g(\cdot)$ is uniformly continuous on $[0, 1]$. Thus, for all $\varepsilon > 0$, there exists $\delta > 0$, such that

$$|g(x) - g(x')| < \varepsilon, \ \forall x, x' \in [0, 1] \text{ with } |x - x'| < \delta.$$  \hfill (2.2)

Since $X \in \text{RV}_{-\alpha}$ with $\mathbb{P}(X > t) = t^{-\alpha}L(t), \ t > 0, \ \alpha > 0$, where $L$ is a slowly varying function, it follows from Lemma 2.1 that

$$\frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} = \frac{(wt)^{-\alpha}L(wt)}{t^{-\alpha}L(t)} \rightarrow w^{-\alpha}, \text{ as } t \rightarrow \infty,$$ \hfill (2.3)

uniformly on $[1, \infty)$. The uniform convergence of (2.3) implies that for the selected $\delta > 0$ in (2.2), there exists an $N > 0$, whenever $t > N$,

$$\left| \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} - w^{-\alpha} \right| < \delta, \ \forall w \in [1, \infty).$$

Since both $\frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)}$ and $w^{-\alpha}$ are in $[0, 1]$ for $w \geq 1$, we have from (2.2) that whenever $t > N$,

$$\left| g\left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) - g(w^{-\alpha}) \right| < \varepsilon, \ \forall w \in [1, \infty).$$

The uniform convergence holds.

Applying Proposition 2.2, we obtain that as $p \rightarrow 1$,

$$H_g(X| X > \text{VaR}_p(X)) \approx \left( 1 + \int_{1}^{\infty} g(w^{-\alpha})dw \right) \text{VaR}_p(X),$$

for any continuous distortion function $g$. The assumption of continuity can be weakened by using the following deeper analysis.

**Theorem 2.3.** If a loss random variable $X \in \text{RV}_{-\alpha}$, and $g(\cdot)$ is any distortion function with $\int_{1}^{\infty} g(w^{-\alpha+\delta})dw < \infty$ for some $0 < \delta < \alpha$, then for any $b \geq 1$,

$$\lim_{t \rightarrow \infty} \int_{b}^{\infty} g\left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) dw = \int_{b}^{\infty} g(w^{-\alpha})dw.$$  

**Proof.** Since $g$ is nondecreasing, $\int_{b}^{\infty} g(w^{-\alpha})dw \leq \int_{b}^{\infty} g(w^{-\alpha+\delta})dw < \infty$ for any $b \geq 1$, then for any $\varepsilon > 0$, there exists an $N_1$ such that

$$\int_{N_1}^{\infty} g(w^{-\alpha})dw < \frac{\varepsilon}{4}. \hfill (2.4)$$

Since $X \in \text{RV}_{-\alpha}$, we have

$$\frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} = \frac{(wt)^{-\alpha}L(wt)}{t^{-\alpha}L(t)} = \frac{w^{-\alpha}L(wt)}{L(t)},$$
where $L$ is a slowly varying function. By the representation theorem of slowly varying functions (see, e.g., Theorem 1.3.1 in [2] for detail), we know that

$$L(x) = c(x) \exp \left\{ \int_a^x \varepsilon(u) du / u \right\}, \quad x \geq a$$

for some $a > 0$, where $c(\cdot)$ is measurable and bounded, and $c(x) \to c > 0$ as $x \to \infty$, and $\varepsilon(x) \to 0$ as $x \to \infty$. Therefore,

$$\frac{P(X > wt)}{P(X > t)} = \frac{w^{-\alpha}L(wt)}{L(t)} = w^{-\alpha} \frac{c(wt)}{c(t)} \exp \left\{ \int_t^{wt} \varepsilon(u) du / u \right\}, \quad \text{for } w \geq b.$$  

For given $\delta > 0$, we have that for sufficiently large $t$, $-\delta < \varepsilon(t) < \delta$, so that

$$\int_t^{wt} \varepsilon(u) du / u \leq \delta \int_t^{wt} du / u = \delta \log w.$$  

Since $c(wt)/c(t)$ is asymptotically bounded by, say, $c^* \geq 1$, as $t \to \infty$, there exists a $t_0$ such that whenever $t > t_0$,

$$\frac{P(X > wt)}{P(X > t)} \leq w^{-\alpha} c^* \exp \{\delta \log w\} = c^* w^{-\alpha + \delta}, \quad \text{for } w \geq b \geq 1. \quad (2.5)$$

Since $g(\cdot)$ is nondecreasing, the upper bound in (2.5) implies that there exists an $N_2$ such that whenever $t > t_0$,

$$\int_{N_2}^\infty g \left( \frac{P(X > wt)}{P(X > t)} \right) dw \leq \int_{N_2}^\infty g(c^* w^{-\alpha + \delta}) dw < \frac{\varepsilon}{4}. \quad (2.6)$$

Let $N = \max\{N_1, N_2\}$. Since $g(\cdot)$ is non-decreasing and bounded, the set of discontinuity points of $g(\cdot)$ is at most countable (Froda’s theorem) and thus has Lebesgue measure zero. That is,

$$\lim_{t \to \infty} g \left( \frac{P(X > wt)}{P(X > t)} \right) = g(w^{-\alpha}), \quad (2.7)$$

almost everywhere on $[1, N]$. By the third Littlewood’s principle (see, e.g., Proposition 24, [15], page 73), there exists a set $A \subseteq [1, N]$ such that $\mu(A) \leq \epsilon/8$, where $\mu(\cdot)$ denotes the Lebesgue measure, and the convergence in (2.7) is uniform on $[1, N] \setminus A$. That is, there is a $t_1 \geq t_0$, such that whenever $t > t_1$,

$$\left| g \left( \frac{P(X > wt)}{P(X > t)} \right) - g(w^{-\alpha}) \right| \leq \frac{\epsilon}{4(N - b)}, \quad \forall \ w \in [1, N] \setminus A.$$  

8
which, together with the fact that $g$ is bounded by 1, imply that, whenever $t > t_1 \geq t_0$,

$$\left| \int_b^N g \left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) dw - \int_b^N g(w^{-\alpha}) dw \right| \leq \int_A g \left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) - g(w^{-\alpha}) \right| dw
\leq 2 \int_A dw + \frac{\varepsilon}{4} \leq \frac{\varepsilon}{2}. \quad (2.8)$$

Combining (2.4), (2.6) and (2.8), we have that for given $\varepsilon > 0$, whenever $t > t_1$,

$$\left| \int_b^\infty g \left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) dw - \int_b^\infty g(w^{-\alpha}) dw \right|
\leq \left| \int_b^N g \left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) dw - \int_b^N g(w^{-\alpha}) dw \right|
+ \left| \int_\infty^\infty g \left( \frac{\mathbb{P}(X > wt)}{\mathbb{P}(X > t)} \right) dw \right|
+ \left| \int_N^\infty g(w^{-\alpha}) dw \right|
\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \quad (2.9)$$

The desired limit follows. \hfill \Box

Since a distortion function $g(\cdot)$ is uniformly continuous on $[0, 1]$ except perhaps for a subset with small Lebesgue measure, the detailed representation of regular variation has to be used to estimate the tails of the involved integrals. Applying Theorem 2.3 to (2.1) yields our main result.

**Theorem 2.4.** If $X \in \text{RV}_{-\alpha}$ and $g(\cdot)$ is any distortion function with $\int_1^\infty g(w^{-\alpha+\delta}) dw < \infty$ for some $0 < \delta < \alpha$, then,

$$\lim_{p \to 1} \frac{H_g(X \mid X > \text{VaR}_p(X))}{\text{VaR}_p(X)} = 1 + \int_1^\infty g(w^{-\alpha}) dw.$$ 

**Remark 2.5.** 1. If $g(\cdot)$ is continuous, then by Proposition 2.2, $\int_1^\infty g(w^{-\alpha}) dw$ must be finite in order to use our estimate for tail distortion. For example, if $g(t) \approx t^k$ as $t \to 0$, where $k > 1/\alpha$, then $\int_1^\infty g(w^{-\alpha}) dw < \infty$.

2. In the case that discontinuity exists, by Theorem 2.4, $\int_1^\infty g(w^{-\alpha+\delta}) dw$ must be finite in order to use our estimate for tail distortion. This is slightly stronger than requiring that $\int_1^\infty g(w^{-\alpha}) dw < \infty$. For example, if $g(t) \approx t^k$ as $t \to 0$, where $k > 1/(\alpha - \delta)$ ($> 1/\alpha$), then $\int_1^\infty g(w^{-\alpha+\delta}) dw < \infty$. 

9
Note that if $g = 1$ and $\alpha > 1$, then the asymptotic relation in Theorem 2.4 reduces to (1.9). It is also easy to see that the ratio of $H_g(X \mid X > \text{VaR}_p(X))$ over $\text{VaR}_p(X)$ is asymptotically decreasing to 1 as tail index $\alpha$ increases to $\infty$.

**Example 2.6.** (Proportional Hazards Distortion)

The proportional hazards transform is obtained by constraining the expected integrated hazard rate under the deformed distribution. This distortion has the following form:

$$g(x) = x^k, \quad 0 \leq x \leq 1, \quad k \geq 0. \tag{2.9}$$

This distortion function is concave if and only if $k \leq 1$, and convex if and only if $k \geq 1$. Hence, the distortion risk measure is coherent if and only if $k \leq 1$. This distortion risk measure has been extensively studied in insurance applications (see [19], [20] and [22]).

For a loss random variable $X \in \text{RV}_{-\alpha}$, where $\alpha > 1/k$, we have, as $t \to \infty$,

$$H_g(X \mid X > t) \approx 1 + \int_1^\infty g(w^{-\alpha})dw = 1 + \int_1^\infty w^{-\alpha k}dw = 1 + \frac{1}{\alpha k - 1} = \frac{\alpha k}{\alpha k - 1}.$$

If we choose $t = \text{VaR}_p(X)$, then we have that as $p \to 1$,

$$H_g(X \mid X > \text{VaR}_p(X)) \approx \frac{\alpha k}{\alpha k - 1} \text{VaR}_p(X). \tag{2.10}$$

Note that if $0 < \alpha \leq 1$, then TCE of $X$ does not exist. In contrast, the tail distortion risk of $X$ with distortion function with $k > 1/\alpha$ can still be estimated via (2.10).

**Example 2.7.** (Exponential Distortion)

The exponential distortion is simply the cumulative distribution function of an exponential random variable constrained to the unit interval. Here we have:

$$g(x) = \frac{1 - e^{-x/c}}{1 - e^{-1/c}}, \quad 0 \leq x \leq 1. \tag{2.11}$$

The exponential distortion function is always concave, therefore, the corresponding distortion risk measure is coherent.

Consider an exponential distortion function $g(x) = \frac{1 - e^{-x}}{1 - e^{-1}}, \quad x \in [0, 1]$ and a loss random variable $X \in \text{RV}_{-\alpha}$, $\alpha > 1$. As $t = \text{VaR}_p(X) \to \infty$, we have

$$H_g(X \mid X > t) \approx 1 + \int_1^\infty g(w^{-\alpha})dw = 1 + \int_1^\infty \frac{1 - e^{-1/w^\alpha}}{1 - e^{-1}}dw.$$
For various tail index values, we tabulate the asymptotics as follows:

| $\alpha$ | $\int_1^\infty \frac{1-e^{-t/w}}{1-e^{-1}}dw$ | $\lim_{t \to \infty} H_g(X|X>t)$ |
|----------|---------------------------------|------------------|
| 1.5      | 2.832                           | 3.832            |
| 2        | 1.363                           | 2.363            |
| 2.5      | 0.891                           | 1.891            |
| 3        | 0.661                           | 1.661            |
| 3.5      | 0.524                           | 1.524            |
| 4        | 0.435                           | 1.435            |

3 Tail Distortion for Location, Scale and Shape Invariant Families

In this section, we discuss three distortion function families, resulting the distorted distributions that are invariant, respectively, under location, scale and shape transformations. In each case, we calculate the tail distortion via Theorem 2.4.

3.1 Location invariant distortion

Let $\mathcal{F}_l$ denote the family of distributions on $\mathbb{R}$ that are invariant under location transforms; that is, if the distribution of a random variable $W$ is in $\mathcal{F}_l$, then the distribution of $W - \lambda$ is also in $\mathcal{F}_l$ for any $\lambda \in \mathbb{R}$. Define

$$g_{F,\lambda}(x) = F(F^{-1}(x) + \lambda), \quad x \in [0,1],$$  

(3.1)

where $\lambda \in \mathbb{R}$. Obviously, (3.1) is a well defined distortion function, and if $F$ is strictly increasing, then $F \in \mathcal{F}_l$ implies that $g_{F,\lambda}(F(x)) \in \mathcal{F}_l$ for any $\lambda$.

The examples of (3.1) include the following Wang transform, introduced in [21] to develop a universal pricing method:

$$F^*(x) = \Phi[\Phi^{-1}(F(x)) + \lambda],$$  

(3.2)

where $\Phi$ is the standard normal cumulative distribution. The parameter $-\lambda$ is called the market price of risk, reflecting the level of systematic risk, and $F(x)$ is the cumulative distribution of a financial asset value $X$. For a liability with loss variable $X$, the Wang transform has an equivalent representation.

$$F^*(x) = \Phi[\Phi^{-1}(1 - F(x)) + \lambda],$$  

(3.3)
which can be used in the set-up of Definitions 1.1 and 1.4. Figure 1 shows the Wang transform with various parameters. The distortion function (3.1) and its dynamic extension were proposed in [7] under the guise of Esscher-Girsanov transform. The connection between the Esscher-Girsanov transform and the Wang transform is highlighted in [11].

**Example 3.1.** Suppose a loss random variable \( X \in \text{RV}_{-\alpha}, \; \alpha > 1 \). For the distortion function \( g(x) = \Phi[\Phi^{-1}(x) + \lambda], \; x \in [0, 1] \), as \( t \to \infty \), we have

\[
H_g(X \mid X > t) \approx 1 + \int_1^\infty \Phi[\Phi^{-1}(w^{-\alpha}) + \lambda]dw. \quad (3.4)
\]

For various values of tail index \( \alpha \) and market price \( -\lambda \), we tabulate the corresponding tail estimates as follows.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \lim_{t \to \infty} \frac{H_g(X \mid X &gt; t)}{t}, \lambda = -0.7 )</th>
<th>( \lim_{t \to \infty} \frac{H_g(X \mid X &gt; t)}{t}, \lambda = 0.25 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>1.620</td>
<td>4.199</td>
</tr>
<tr>
<td>2</td>
<td>1.379</td>
<td>2.429</td>
</tr>
<tr>
<td>2.5</td>
<td>1.274</td>
<td>1.913</td>
</tr>
<tr>
<td>3</td>
<td>1.215</td>
<td>1.670</td>
</tr>
<tr>
<td>3.5</td>
<td>1.177</td>
<td>1.528</td>
</tr>
<tr>
<td>4</td>
<td>1.150</td>
<td>1.436</td>
</tr>
</tbody>
</table>

Table 1: Tail Estimations for Wang Transform
3.2 Scale invariant distortion

Let \( F_s \) denote the distribution family invariant under scale transforms; that is, if the distribution of \( W \) is in \( F_s \), then the distribution of \( W/\sigma \) is also in \( F_s \) for any \( \sigma > 0 \). Define

\[
g_{F,\sigma}(x) = F(F^{-1}(x) \cdot \sigma), \quad x \in [0, 1],
\]

(3.5)

where \( \sigma \in \mathbb{R}_+ \). The function in (3.5) is also a well defined distortion function, and if \( F \) is strictly increasing, then \( F \in F_s \) implies that \( g_{F,\sigma}(F(x)) \in F_s \) for any \( \sigma > 0 \). For example, the family of the normal distributions are scale invariant. Figure 1 shows scale invariant distortions for normal distributions.

Example 3.2. Suppose a loss random variable \( X \in \text{RV}_{-\alpha}, \alpha > 1 \). For the distortion function \( g(x) = \Phi[\Phi^{-1}(x) \cdot \sigma], \quad x \in [0, 1] \), as \( t \to \infty \), we have

\[
\frac{H_g(X|X > t)}{t} \approx 1 + \int_{1}^{\infty} \Phi[\Phi^{-1}(w^-\alpha) \cdot \sigma]dw.
\]

(3.6)

For various tail index values, Table 2 shows tail estimations for scale invariant distortion with different parameters.

| \( \alpha \) | \( \lim_{t \to \infty} \frac{H_g(X|X > t)}{t}, \sigma = 2 \) | \( \lim_{t \to \infty} \frac{H_g(X|X > t)}{t}, \sigma = 0.9 \) |
|---------|-----------------|-----------------|
| 1.5     | 1.751           | 4.553           |
| 2.5     | 1.383           | 1.808           |
| 4.5     | 1.193           | 1.321           |
| 5       | 1.171           | 1.279           |
| 6       | 1.140           | 1.221           |
| 7       | 1.119           | 1.183           |

Table 2: Tail Estimations for Scale Invariant Distortion

3.3 Power distortion

For any nonnegative, regularly varying loss variable \( X \in \text{RV}_{-\alpha} \), it follows from Proposition 2.6 in [14] that \( X^{\frac{1}{k}} \in \text{RV}_{-\alpha k} \). Define, for the distribution \( F \) of a nonnegative random variable,

\[
g_{F,k}(x) = F((F^{-1}(x))^k), \quad x \in [0, 1],
\]

(3.7)

where \( k > 0 \). It is easy to see that (3.7) is a well defined distortion function, and if \( F \) is strictly decreasing and has a regularly varying right tail with tail index \( \alpha \), then \( g_{F,k}(F(x)) \) also has a regularly varying right tail with tail index \( k\alpha \).
If we choose $F$ as the Pareto distribution, Figure 2 shows power distortions with Pareto distributions for various $k$’s and different shape parameters (tail indexes). Figure 3 shows tail estimations for power distortions with Pareto distributions ($k = 2$). It is worth mentioning that the power distortion with Pareto distribution with shape parameter $\alpha = 2$ resembles the Pareto distortion discussed in [5, 18].

**Example 3.3.** As in [21], consider a ground-up liability risk $X$ with a Pareto survival (severity) distribution

$$S(x) = \left(\frac{2000}{2000 + x}\right)^{1.2}, \text{ for } x > 0.$$ 

To compare the risk loading, we apply the power distortion $g_{S,k}(x) = S((S^{-1}(x))^k)$ with various values for $k$ and we calculate the risk loading with and without distortion. The resulting summary of the expectations of the excess of loss amount is provided in Table 3.

The risk adjustment and loading percentage increase as the power of distortion function decreases. This can be explained by the fact that there is greater uncertainty in the tails of

<table>
<thead>
<tr>
<th>Tail Conditional Expectation</th>
<th>$k = 0.99$</th>
<th>$k = 0.97$</th>
<th>$k = 0.95$</th>
<th>$k = 0.93$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without distortion $\mathbb{E}(X \mid X &gt; \text{VaR}_{0.95}(X))$</td>
<td>$132882$</td>
<td>$132882$</td>
<td>$132882$</td>
<td>$132882$</td>
</tr>
<tr>
<td>With distortion $H_g(X \mid X &gt; \text{VaR}_{0.95}(X))$</td>
<td>$149005$</td>
<td>$191704$</td>
<td>$255532$</td>
<td>$355924$</td>
</tr>
<tr>
<td>Risk Adjustment</td>
<td>$16123$</td>
<td>$58822$</td>
<td>$122650$</td>
<td>$223042$</td>
</tr>
<tr>
<td>Loading Percentage</td>
<td>12.1%</td>
<td>44.3%</td>
<td>92.3%</td>
<td>167.8%</td>
</tr>
</tbody>
</table>

Table 3: Summary of Risk Adjustments
the distributions as tail index decreases (that is, the tail becomes heavier).

4 Conclusion

In this paper, we have introduced the tail distortion risk for extreme loss, and derived the explicit, asymptotic expression of the tail distortion for losses that have regularly varying right tails. In contrast to distortion risks, the tail distortion risk measures for heavy tailed losses admit asymptotically linear relations with VaR with the proportionality constant depending only on the distortion function and tail index. The well-known asymptotic relation between TCE and VaR serves as a special case of the result we derived.

We conclude this paper by illustrating an application of our results to insurance premium pricing. The calculation of premium is usually affected by various elements: pure risk premium, risk loading, administrative expenses, and loading for investment risk, credit risk, operational risk, etc. We consider here a simplified premium pricing model that consists of the pure risk premium and risk loading only. Let $X$ denote the total claim incurred for one insurance portfolio, having a regularly varying loss distribution with tail index $\alpha$. The pure risk premium is the mean $E(X)$ of loss $X$, and the risk loading depends on the excess of a risk measurement over the mean loss. If we are interested in the tail of the underlying loss distribution and adopt a tail distortion measure $H_g(X|X > \text{VaR}_p(X))$ for $p$ near 1, then the risk loading depends on the excess $H_g(X|X > \text{VaR}_p(X)) - E(X)$ of
the tail risk measurement over the pure risk premium. That is, if we hold the risk capital $H_g(X | X > \text{VaR}_p(X)) \geq \text{VaR}_p(X)$, we know we wouldn’t have ruin with probability $p$. If an insurance company receives the excess $H_g(X | X > \text{VaR}_p(X)) - E(X)$ from investors and invests the capital at the risk-free rate $r_0$ (e.g., in government bonds), the company needs to pay the investors at a higher rate $r > r_0$, because their investment is exposed to risk. Thus, the premium paid by the policyholder in this simple pricing model is the sum of the pure risk premium and risk loading:

$$E(X) + (r - r_0)(H_g(X | X > \text{VaR}_p(X)) - E(X)) = (1 - r + r_0)E(X) + (r - r_0)H_g(X | X > \text{VaR}_p(X)).$$

Since $p$ is close to 1, Theorem 2.4 implies that the premium is approximately equal to

$$(1 - r + r_0)E(X) + (r - r_0)\left(1 + \int_1^{\infty} g(w^{-\alpha})dw\right)\text{VaR}_p(X).$$

The loss distribution and model parameters can be fitted from claim data, and the calculation of the premium is then straightforward.

**Acknowledgments:** The authors would like to sincerely thank a referee for his/her insightful comments, which led to significant improvements of the results and presentation of this paper.

**References**


