A primer on generalized weighted risk functionals

Nawaf Mohammed *[1, 3], Edward Furman[1, 3], and Jianxi Su[2, 3]

[1]Department of Mathematics and Statistics, York University, Toronto, ON M3J 1P3, Canada.

[2]Department of Statistics, Purdue University, West Lafayette, IN 47906, USA

[3]Risk and Insurance Studies Centre, 4700 Keele St, Toronto, ON M3J 1P3, Canada

Abstract

Weighted risk functionals have been well-studied to date. Indeed, this versatile class of risk functionals has enjoyed a variety of applications in risk management and insurance and has been generalized along a number of directions. Namely, some exemplary applications of the class of weighted risk functionals in risk management and insurance include pricing, valuation, risk measurement, risk capital allocation, and risk sharing, whereas existing variations of weighted risk functionals allow for such enhancements as the multivariate probability weighting and the augmentation of utility functions.

In this paper, we propose a class of generalized weighted risk functionals that incorporates the possibility of arbitrary loss aggregations. To this end, we introduce the notion of aggregation function in the context of the mentioned weighted risk functionals. Then we delineate the ways in which distinct orders on the weight functions and on the aggregation functions impact the orders of the generalized weighted risk functionals introduced herein. We conclude with several observations that facilitate applications of the generalized weighted risk functionals.

Key words and phrases: Weighted distribution; weighted premium calculation premium; weighted risk capital allocation; aggregation function; regression function

JEL Classification: C92, G32

*Corresponding author; postal address: 4700 Keele St, Toronto, ON M3J 1P3, Canada; email: nawaf@yorku.ca

1 Introduction

Consider non-negative random variables (RVs) $X,\ X_1,\dots,X_n,\ n\in\mathbb{N}$, representing (insurance) losses, and let $\mathcal{X}$ denote a collection of such losses. For a Borel-measurable non-negative – and as a rule non-decreasing – 'weight' function $x\mapsto w(x),\ x\in[0,\ +\infty)$, the functionals $H_w:\mathcal{X}\to[0,\ +\infty)\cup\{+\infty\}$, such that the ratio of expectations below is well-defined and finite

$$H_w(X)=\frac{\mathbb{E}[Xw(X)]}{\mathbb{E}[w(X)]}, \tag{1}$$

are often called 'weighted' risk measures; also called actuarial premium calculation principles, if the bound $H_w(X)\geq\mathbb{E}[X]$ holds for those RVs $X\in\mathcal{X}$ that have finite means (e.g., Sendov et al., 2011), that is the non-negative loading property is satisfied. Most recently. the class of weighted functionals, H_w, has been connected to a theory of stress-testing, in which case weight functions play the role of 'stressing' mechanisms (e.g., Millossovich et al., 2021). In what follows, H_w is referred to as the weighted risk functional(s) to recognize the manifold of existing applications across risk management and insurance.

In actuarial science, weighted risk functionals, H_w, were introduced by Furman and Zitikis (2007, 2008a) as a unifying class of risk functionals that comprises, e.g., the Value-at-Risk and the Conditional Tail Expectation risk measures, Esscher's, Kamps', and – under certain conditions – the distorted premiums, among other popular risk measures and actuarial premiums. We refer to, e.g., Choo and de Jong (2009); Kaluszka and Krzeszowiec (2012) and a more recent Castano-Martinez et al. (2020) and references therein, for examples of works that explore properties of the class of weighted risk functionals.

Generalizations of (1) have been developed in several directions, with the arguably simplest and most-popular of these directions having led to the rise of the notion of weighted risk capital allocations, put forward in Furman and Zitikis (2008b). More specifically, let $S=X_1+\cdots+X_n$ denote the aggregate loss RV, then the functionals $A_w(X_i,S):\mathcal{X}\times\mathcal{X}\to[0,\ +\infty)\cup\{+\infty\}$, such that the ratio of expectations below is well-defined and finite

$$A_w(X_i, S) = \frac{\mathbb{E}[X_i w(S)]}{\mathbb{E}[w(S)]}, \quad i \in \{1, \dots, n\}, \tag{2}$$

are called weighted risk capital allocations (e.g., Dhaene et al., 2012; Guo et al., 2018, for details).

An alternative generalization of (1), which is referred to as a generalized weighted risk measure or premium in Furman and Zitikis (2009), is obtained by considering the class of functionals $H_{v,w} : \mathcal{X} \to [0, +\infty) \cup \{+\infty\}$, such that

$$H_{v,w}(X) = \frac{\mathbb{E}[v(X)w(X)]}{\mathbb{E}[w(X)]}, \tag{3}$$

where v, w are non-negative and Borel-measurable functions, $\mathbb{E}[v(X)w(X)] \in \mathbb{R}_+$, and $\mathbb{E}[w(X)] \in \mathbb{R}_+$ (e.g., Richards and Uhler, 2019, for a study of the monotonicity of the class of generalized weighted risk functionals).

Yet another generalization of weighted risk functionals (1) was considered in Millossovich et al. (2021); Porth et al. (2014); Zhu et al. (2019) (also, Furman and Zitikis, 2007, for an earlier note in this respect). This generalization hinges on the assumption that the weight function $w(\cdot) \geq 0$ – non-decreasing in each variable and Borel-measurable – operates on vectors of loss RVs, that is $w : [0, \infty)^n \to [0, \infty)$. Clearly, if the weight function is chosen to be the 'simple sum' aggregation function, that is $w(\boldsymbol{x}) = x_1 + \cdots + x_n,\ \boldsymbol{x} = (x_1, \dots x_n) \in [0, +\infty)^n$, then functional (2) is recovered. Zhu et al. (2019) focus on linear and log-linear combinations of rate-making factors as the weight functions of interest and derive properties of what they call 'multivariate' weighted premiums (for various weight functions that arise in the context of a multivariate stress-testing theory, we refer to Millossovich et al. (2021).

Speaking generally, aggregate financial positions are not simple sums of loss RVs (e.g., (e.g., Chapter 5 in Jaworski et al., 2010). Namely, let the function $g : [0, +\infty)^n \to [0, +\infty)$ be non-decreasing in each variable, Borel-measurable, and, for $\boldsymbol{x} \in [0, +\infty)^n$, satisfy the boundary conditions

$$0 \leq \inf_{\boldsymbol{x} \in [0,+\infty)^n} g(\boldsymbol{x}) < \infty, \quad \text{and} \quad 0 < \sup_{\boldsymbol{x} \in [0,+\infty)^n} g(\boldsymbol{x}) \leq +\infty,$$

that is the function $\boldsymbol{x} \mapsto g(\boldsymbol{x})$ is a general aggregation function (e.g., Grabisch et al., 2009), and let $S_g = g(\boldsymbol{X})$, $\boldsymbol{X} = (X_1, \ldots, X_n) \in \mathcal{X}^n$, denote the g-aggregate loss RV. Examples of aggregation functions are the already-mentioned (e.g., Zhu et al., 2019) simple sum aggregation function $g(\boldsymbol{x}) = \sum_{i=1}^n x_i$ and the exponential aggregation function $g(\boldsymbol{x}) = \sum_{i=1}^n e^{x_i}$. Other examples of aggregation functions are, e.g.,

- the maximum aggregation function – also, the largest order statistic – $g(\boldsymbol{x}) = \max(x_1, \ldots, x_n)$;
- the minimum aggregation function – also, the smallest order statistic – $g(\boldsymbol{x}) = \min(x_1, \ldots, x_n)$;
- the product aggregation function $g(\boldsymbol{x}) = x_1 \times \cdots \times x_n$;
- the log-sum-exp aggregation function $g(\boldsymbol{x}) = \log(e^{x_1} + \cdots + e^{x_n})$;
- the p-norm aggregation function $g(\boldsymbol{x}) = \left(x_1^p + \cdots + x_n^p\right)^{1/p}$, where $p \in \mathbb{R}_+$.

In this paper, we work with the class of g-aggregation functions, such that the projection onto the i-th variable, P_i, $i \in \{1, \ldots, n\}$, equals that variable; namely, we require $P_i[g(\boldsymbol{x})] = g(x_i) = x_i$. This additional condition, which by passing implies that the class of weight functions and the class of aggregation functions do not generally agree, is natural as an aggregation of a singleton is not really an aggregation. Keeping the above in mind, in this paper we work with the following generalized weighted risk functionals

$$H_w(S_g) = \frac{\mathbb{E}\left[S_g \times w \circ S_g\right]}{\mathbb{E}\left[w \circ S_g\right]} \tag{4}$$

and, for $i \in \{1, \ldots, n\}$,

$$A_w(X_i, S_g) = \frac{\mathbb{E}\left[X_i \times w \circ S_g\right]}{\mathbb{E}\left[w \circ S_g\right]}; \tag{5}$$

unless stated otherwise, we assume in the sequel that the ratios of expectations above are finite and well-defined. Clearly, if the g-aggregation function is the simple sum aggregation, then weighted risk functionals (4) and (5) reduce to the original ones of Furman and Zitikis (2008a,b). Table 1 presents some weighted functionals for the popular choices of weight functions.

Name	$w(y)$	$H_w(S_g)$	$A_w(X_i, S_g)$
Net	const	$\mathbb{E}[S_g]$	$\mathbb{E}[X_i]$
Modified variance	y	$\mathbb{E}[S_g] + \frac{\mathrm{Var}(S_g)}{\mathbb{E}[S_g]}$	$\mathbb{E}[X_i] + \frac{\mathrm{Cov}(X_i,\, S_g)}{\mathbb{E}[S_g]}$
Size-biased	y^t	$\frac{\mathbb{E}\big[{S_g}^{1+t}\big]}{\mathbb{E}\big[{S_g}^t\big]}$	$\frac{\mathbb{E}\big[X_i \times {S_g}^t\big]}{\mathbb{E}\big[{S_g}^t\big]}$
Esscher	$e^{t\,y}$	$\frac{\mathbb{E}\big[S_g \times \exp\left(t\,S_g\right)\big]}{\mathbb{E}\big[\exp\Big(t\,S_g\Big)\big]}$	$\frac{\mathbb{E}\big[X_i \times \exp\left(t\,S_g\right)\big]}{\mathbb{E}\big[\exp\left(t\,S_g\right)\big]}$
Aumann-Shapley	$e^{t\,F(y)}$	$\frac{\mathbb{E}\big[S_g \times \exp\left(t\,F(S_g)\right)\big]}{\mathbb{E}\big[\exp\left(t\,F(S_g)\right)\big]}$	$\frac{\mathbb{E}\big[X_i \times \exp\left(t\,F(S_g)\right)\big]}{\mathbb{E}\big[\exp\left(t\,F(S_g)\right)\big]}$
Kamps	$1 - e^{-t\,y}$	$\frac{\mathbb{E}\big[S_g \times \left(1 - e^{-tS_g}\right)\big]}{\mathbb{E}\big[1 - e^{-t\,S_g}\big]}$	$\frac{\mathbb{E}\big[X_i \times \left(1 - e^{-t\,S_g}\right)\big]}{\mathbb{E}\big[1 - e^{-t\,S_g}\big]}$
Conditional tail expectation	$\mathbb{1}\{y \geq y_q\}$	$\mathbb{E}\big[S_g \vert S_g \geq s_q\big]$	$\mathbb{E}\big[X_i \vert S_g \geq s_q\big]$
Modified tail variance	$y\,\mathbb{1}\{y \geq y_q\}$	$\mathbb{E}\big[S_g \vert S_g \geq s_q\big] + \frac{\mathrm{Var}\big(S_g \vert S_g \geq s_q\big)}{\mathbb{E}\big[S_g \vert S_g \geq s_q\big]}$	$\mathbb{E}\big[X_i \vert S_g \geq s_q\big] + \frac{\mathrm{Cov}\big(X_i,\, S_g \vert S_g \geq s_q\big)}{\mathbb{E}\big[S_g \vert S_g \geq s_q\big]}$
Distorted	$h^{'}\left(\overline{F}(y)\right)$	$\mathbb{E}\big[S_g \times h^{'}\left(\overline{F}(S_g)\right)\big]$	$\mathbb{E}\big[X_i \times h^{'}\left(\overline{F}(S_g)\right)\big]$
Proportional hazard	$q\left(\overline{F}(y)\right)^{q-1}$	$q\,\mathbb{E}\big[S_g \times \left(\overline{F}(S_g)\right)^{q-1}\big]$	$q\,\mathbb{E}\big[X_i \times \left(\overline{F}(S_g)\right)^{q-1}\big]$

Table 1: Examples of weight function w accompanied with their associate risk measures and allocations. Within the table, we let $S_g = g(\boldsymbol{X})$, F and $\overline{F}$ denote the cumulative and decumulative distribution functions of S_g, respectively. Moreover, the distortion function $h : [0,1] \mapsto [0,1]$ is non-decreasing such that $h(0) = 0$ and $h(1) = 1$. Lastly, y_q, $q \in (0,1)$, is a threshold that is dependent on an a priori probability level q where s_q is the Value-at-Risk at that level.

The rest of this paper is devoted to the study of various properties of functionals (4) and (5). More specifically in Section 2, we investigate bounds for the pairs of weighted risk functionals $H_w(S_{g_1})$ and $H_w(S_{g_2})$ as well as $A_w(X_i, S_{g_1})$ and $A_w(X_i, S_{g_2}),\ i \in \{1, \ldots, n\}$, where the weight function, w, is fixed and two distinct g-aggregation functions, g_1 and g_2, are considered. Notably, by selecting the g-aggregation functions, g_1 and g_2, such that $g_2 = \xi \circ g_1$ with the appropriately chosen non-decreasing and Borel-measurable function $\xi : [0, +\infty) \to [0, +\infty)$, the results in Section 2 help compare the riskiness of aggregate losses subject to coverage modifications. Then in Section 3, we repeat the exercise by comparing weighted risk functionals $H_{w_1}(S_g)$ and $H_{w_2}(S_g)$ as well as $A_{w_1}(X_i, S_g)$ and $A_{w_2}(X_i, S_g),\ i \in \{1, \ldots, n\}$, which this time share the same aggregation function, g, but have different weight functions, w_1 and w_2. Not surprisingly, a departure from the simple sum aggregation function results in a significant layer of complexity both when studying properties of generalized weighted risk functionals (4) and (5) and when evaluating them. In Section 4, we characterise those loss RVs, for which – irrespective of the choice of the g-aggregation function and the weight function – risk functional (5) is either trivially obtained from risk functional (4) or equals a constant (e.g., Guan et al., 2021, for a related discussion). Section 5 concludes the paper.

2 Orders based on different aggregation functions but the same weight function

In what follows, we fix an atomless probability space and denote by $\mathcal{X}$ and $\mathcal{X}^n$ the set of all non-negative RVs and the set of all non-negative random vectors $\boldsymbol{X} = (X_1, \ldots, X_n)$; in both cases these are interpreted as losses in a portfolio of losses, $\mathcal{N} = \{1, \ldots, n\},\ n \in \mathbb{N}$. The cumulative distribution function and the decumulative distribution function of the RVs $X \in \mathcal{X}$ and $\boldsymbol{X} \in \mathcal{X}^n$ are denoted by $F_X(x) = \mathbb{P}(X \le x),\ \overline{F}_X(x) = 1 - F_X(x)$ and $F_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{X} \le \boldsymbol{x}),\ \overline{F}_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{X} > \boldsymbol{x})$, respectively, for non-negative x and $\boldsymbol{x} = (x_1, \ldots, x_n)$.

It is easy to see that the generalized weighted risk functional as in Equation (5) satisfies the no-unjustified loading property as well as the non-negative loading property (Furman and Zitikis, 2008b) given that the weight function, w, is non-decreasing and the RVs $X_i,\ i \in \mathcal{N}$ and S_g are positively quadrant dependent (PQD) (e.g., Lehmann, 1966, for details). Also, while functionals (5) are fully additive, i.e. $\sum_{i=1}^{n} A_w(X_i, S_g) = H_w(S_g)$, they admit a special form of the no-undercut property only if the g-aggregation function is the simple sum aggregation function ($S_g = S$)(Recall in this respect that the no-undercut property states that stand-alone losses are riskier – require more risk capital – than those losses that are considered a part of a portfolio of losses). The no-undercut property for the class of generalized weighted risk functionals (5) is formulated in the following proposition, which holds due to the Jensen's inequality.

Proposition 1. *If the g-aggregation function is convex, then we have*

$$g\big(A_w(X_1, S_g), \ldots, A_w(X_n, S_g)\big) \leq H_w(S_g) \text{ for all } \boldsymbol{X} \in \mathcal{X}^n \text{ and } S_g = g(\boldsymbol{X}). \tag{6}$$

Proof. Let $\mathbb{E}_{w(S_g)}[\,\cdot\,] = \mathbb{E}[\,\cdot\, w(g(\boldsymbol{X}))]/\mathbb{E}[w(g(\boldsymbol{X}))]$, then we have

$$\begin{aligned} g\big(A_w(X_1, S_g), \ldots, A_w(X_n, S_g)\big) &= g\big(\mathbb{E}_{w(S_g)}[X_1], \ldots, \mathbb{E}_{w(S_g)}[X_n]\big), \\ &\leq \mathbb{E}_{w(S_g)}[g(\boldsymbol{X})] \\ &= H_w(S_g), \end{aligned}$$

which completes the proof. □

Clearly, the p-norm, for $1 \leq p \leq \infty$, as well as the log-sum-exp g-aggregation functions mentioned in Section 1 satisfy the convexity condition in Proposition 1.

Next we turn to the study of how different choices of aggregation functions impact the value of the generalized weighted risk functionals. Two notational conveniences are in place. First, let $w(\cdot)$ be a weight function and let X and Y be two loss RVs in $\mathcal{X}$, all such that the weighted risk functional $A_w(X, Y)$ is well-defined and finite. Then, similarly to the notation in the proof of Proposition 1, $H_w(X, Y) =: \mathbb{E}_{w(Y)}[X]$, where the left-hand side is a w-biased expectation. Similarly, we can write (4) and (5) as $H_w(S_g) = \mathbb{E}_{w(S_g)}[S_g]$ and $A_w(X_i, S_g) = \mathbb{E}_{w(S_g)}[X_i]$, where $S_g = g(\boldsymbol{X})$ and $i \in \mathcal{N}$. Second, let us define the following regression functions, for $\boldsymbol{X} \in \mathcal{X}^n$ and $y \geq 0$,

$$h(y) = \mathbb{E}\big[S_{g_1} \,|\, S_{g_2} = y\big], \quad \text{where } S_{g_j} = g_j(\boldsymbol{X}),\ j = 1, 2$$

and

$$\tilde{h}(y) = \mathbb{E}\big[w(S_{g_2}) \,|\, w(S_{g_1}) = y\big].$$

Theorem 1. *For a weight function w, which is assumed to be strictly increasing, let*

$$H_j = H_w(S_{g_j}), \quad \textit{with } S_{g_j} = g_j(\boldsymbol{X}),\ j = 1, 2,$$

be the weighted risk functionals associated with the aggregation functions g_1 and g_2. The following relationships hold:

$$\textit{If } h(y) \begin{Bmatrix} \geq \\ < \end{Bmatrix} y \textit{ and the function } y \mapsto \frac{y}{\tilde{h}(y)},\ y \in \mathbb{R}_+, \textit{ is } \begin{Bmatrix} \textit{non-decreasing} \\ \textit{decreasing} \end{Bmatrix}, \textit{ then } H_1 \begin{Bmatrix} \geq \\ < \end{Bmatrix} H_2.$$

In particular, $H_1 = H_2$ holds when $h(y) = y$ and the function $y \mapsto y/\tilde{h}(y)$ is constant.

The proof of Theorem 1 is relegated to Appendix Ai.

Remark 1. *The relationship between $y \mapsto h(y)$ and $y \in \mathbb{R}_+$ specified in Theorem 1 compares the order of the realizations of S_{g_1} and S_{g_2}, for a portfolio $\boldsymbol{X} \in \mathcal{X}^n$, in an average sense. Clearly, the order between S_{g_1} and S_{g_2} implies the relationship between $h(y)$ and y. Namely, if $g_1(\boldsymbol{y}) \geq g_2(\boldsymbol{y})$ for all $\boldsymbol{y} \in \mathbb{R}_+^n$, then $h(y) \geq y$ for all $y \in \mathbb{R}_+$. The same argument holds if the inequalities are reversed.*

Remark 2. *In Theorem 1, we assume that the weight function w is strictly increasing. However, this assumption is violated in, e.g., the case of the Excess-of-loss (also, the conditional tail expectation) risk measures and allocations, that is when the weight function is set to be $w(y) = \mathbb{1}(y > d)$ for some $d \geq 0$. In this case, Theorem 1 remains true, but the monotonicity condition of $y \mapsto y/\tilde{h}(y)$ needs to be replaced by that of*

$$y \mapsto \frac{w(y)}{\tilde{h}(y)}, \quad \textit{where } S_{g_j} = g_j(\boldsymbol{X}) \textit{ and } j = 1, 2.$$

We note in passing that the function $y \mapsto y/\tilde{h}(y)$ is non-decreasing if $\tilde{h}(y)$ grows at a slower rate than y does. Note that the monotonicity property of $\tilde{h}$ is related to the dependence structure of the RVs $\widetilde{S}_{g_1} := w(S_{g_1})$ and $\widetilde{S}_{g_2} := w(S_{g_2})$. Specifically, if $\widetilde{S}_{g_2}$ is stochastically non-decreasing in $\widetilde{S}_{g_1}$, i.e.,

$$\mathbb{P}\big(\widetilde{S}_{g_2} > x \,|\, \widetilde{S}_{g_1} = y\big) \text{ is non-decreasing in } y \in \mathbb{R}_+ \text{ for all } x \in \mathbb{R}_+,$$

then the function $y \mapsto \mathbb{E}\big[\widetilde{S}_{g_2} \,|\, \widetilde{S}_{g_1} = y\big]$ is non-decreasing. Whether $\tilde{h}$ increases faster or slower than y depends on the marginal distributions and dependence of the random pair $(\widetilde{S}_{g_1}, \widetilde{S}_{g_2})$, which are stipulated by the choices of g_1, g_2, and w.

The following assertion further clarifies the monotonicity behavior of the function $y \mapsto y/\tilde{h}(y)$ when the dependence of the RVs $\widetilde{S}_{g_1}$ and $\widetilde{S}_{g_2}$ is chosen to be co-monotonic. (The RVs X and Y are said to be co-monotonic if there exist two non-decreasing functions, ξ_1 and ξ_2, and a RV Z, such that $(X, Y) \stackrel{d}{=} (\xi_1(Z), \xi_2(Z))$; here '$\stackrel{d}{=}$' denotes equality in distribution.),

Theorem 2. *For all $\boldsymbol{y} \in \mathbb{R}_+^n$, let $\xi \circ g_1(\boldsymbol{y}) = g_2(\boldsymbol{y})$, where $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ is a non-decreasing function; hence the RVs $\widetilde{S}_{g_1}$ and $\widetilde{S}_{g_2}$ are co-monotonic. Moreover, suppose that the weight function w is diffentiable and log-convex.*

$$\text{If } \xi(y) \begin{Bmatrix} \leq \\ > \end{Bmatrix} y \text{ with } \xi'(y) \begin{Bmatrix} \leq \\ > \end{Bmatrix} 1 \text{ for } y \in \mathbb{R}_+, \text{ then we have } H_1 \begin{Bmatrix} \geq \\ < \end{Bmatrix} H_2.$$

The proof of Theorem 2 is relegated to Appendix Aii.

Among the examples outlined in Table 1, the following weighted risk functionals are associated with a log-convex weight function: Net, Esscher, Aumann-Shapley (when F is convex), distorted (when $h' \circ \overline{F}$ is log-convex), and proportional hazard (when $\overline{F}$ is log-convex).

The mere ordering of the g-aggregation functions is not sufficient in order to have the generalized weighted risk functionals ordered, as becomes evident from the following example.

Example 1. *Suppose that $g_1(\boldsymbol{X}) \sim \mathrm{Pa(II)}(\alpha, \theta)$, that is the RV $g_1(\boldsymbol{X})$ is distributed Pareto of the second kind with the shape parameter $\alpha \in \mathbb{R}_+$, scale parameter $\theta \in \mathbb{R}_+$, and the probability density* *function (PDF) given by*

$$f(x) = \frac{\alpha}{\theta}\left(1+\frac{x}{\theta}\right)^{-(\alpha+1)}, \quad x>0.$$

Further, let $\xi(y) = \max(0, c\,y-d)$*, which represents the risk reduction due to the introduction of a coinsurance factor* $c \in (0,1)$ *and a deductible* $d>0$*. It is straightforward to check that* $\xi(y) \le y$ *and* $\xi'(y) \le 1$*, and thus the first set of assumptions about* ξ *in Theorem 2 is satisfied. Meanwhile, let* $w(y) = y^b$*,* $b>0$*, hence the log-convexity condition on* w *required in Theorem 2 is violated.*

Let $S_g = g_1(\boldsymbol{X})$*, then we have*

$$H_1 = \frac{\mathbb{E}\big[{S_g}^{b+1}\big]}{\mathbb{E}\big[{S_g}^{b}\big]} = \theta\,\frac{1+b}{\alpha-b-1},$$

in which we require $\alpha > 1+b$ *so that the expectations above are well-defined. In a similar fashion, we have*

$$\begin{aligned} H_2 &= \frac{\mathbb{E}\big[\xi(S_g)\times w\circ\xi(S_g)\big]}{\mathbb{E}[w\circ\xi(S_g)]} \\ &= \frac{c^{b+1}\,\mathbb{E}\Big[(S_g-d/c)^{b+1}\,|\,S_g>d/c\Big]}{c^{b}\,\mathbb{E}\Big[(S_g-d/c)^{b}\,|\,S_g>d/c\Big]}. \end{aligned}$$

Note that ${S_g}^*{:=}(S_g-d/c\,|\,S_g>d/c) \sim \mathrm{Pa(II)}\big(\alpha,\, d/c+\theta\big)$*, thus*

$$H_2 = c\,\frac{\mathbb{E}\Big[({S_g}^*)^{b+1}\Big]}{\mathbb{E}\Big[({S_g}^*)^{b}\Big]} = c\,(d/c+\theta)\,\frac{1+b}{\alpha-b-1} = (c\,\theta+d)\,H_1.$$

Set $\theta=1$*, then we obtain* $H_1 \ge H_2$ *if* $c+d\le 1$*, and* $H_1 < H_2$ *if* $c+d>1$*. All in all, this example shows that the order between the aggregation functions* g_1 *and* g_2 *is not sufficient to determine the order between the weighted risk functionals* H_1 *and* H_2*.*

Next we turn to study the impact of the choice of the g-aggregation function on functionals (5). At the outset, let us define, for $S_{g_j} = g_j(\boldsymbol{X})$, $j = 1, 2$, $i \in \mathcal{N}$, and $y \geq 0$, the following regression function

$$\ell_{i,j}(y) = \mathbb{E}\big[w(S_{g_j})|X_i = y\big].$$

Theorem 3. *For a weight function w, let*

$$A_{i,j} = A_w(X_i, S_{g_j}), \ \textit{where } j = 1, 2 \textit{ and } i \in \mathcal{N},$$

be generalized weighted risk functionals á la (5) associated with the aggregation functions g_1 and g_2. The following relationships hold:

$$\textit{If the ratio } y \mapsto \frac{\ell_{i,1}(y)}{\ell_{i,2}(y)} \textit{ is } \left\{ \begin{matrix} \textit{non-decreasing} \\ \textit{decreasing} \end{matrix} \right\} \textit{ on } y \in \mathbb{R}_+, \textit{ then } A_{i,1} \left\{ \begin{matrix} \geq \\ < \end{matrix} \right\} A_{i,2}.$$

In particular, $A_{i,1} = A_{i,2}$ if $y \mapsto \ell_{i,1}(y)/\ell_{i,2}(y)$ is a constant function.

Proof. The result follows from Proposition 3.1 of Furman and Zitikis (2008b). □

Knowing the order between the aggregation functions, g_1 and g_2, may not suffice to determine the monotonicity behavior of the ratio $y \mapsto \ell_{i,1}(y)/\ell_{i,2}(y)$. The following theorem further confirms the critical role that the weight function w plays in shaping the order between the weighted risk functionals $A_{i,1}$ and $A_{i,2}$.

Theorem 4. *Suppose that individual losses within a portfolio have marginal CDFs F_{X_i}, $i \in \mathcal{N}$ and are co-monotonic, namely $X_i \stackrel{d}{=} F_i^{-1}(U)$, where $U \sim \mathrm{Uniform}[0, 1]$. Let the element-wise non-decreasing aggregation functions satisfy $\xi \circ g_1(\boldsymbol{y}) = g_2(\boldsymbol{y})$ for $\boldsymbol{y} \in \mathbb{R}_+^n$, where $\xi : [0, +\infty) \to [0, +\infty)$ is non-decreasing. Further assume that the weight function w is differentiable and log-convex. The following relationships hold*

$$\textit{If } \xi(y) \left\{ \begin{matrix} \leq \\ > \end{matrix} \right\} y \textit{ with } \xi'(y) \left\{ \begin{matrix} \leq \\ > \end{matrix} \right\} 1 \textit{ for } y \in \mathbb{R}_+, \textit{ then we have } A_{i,1} \left\{ \begin{matrix} \geq \\ < \end{matrix} \right\} A_{i,2} \textit{ for } i \in \mathcal{N}.$$

The proof of Theorem 4 is relegated to Appendix Aiii.

In Theorem 4, the log-convexity condition on the weight function w is again minimal, which is reaffirmed in the example below.

Example 2. *Consider two loss RVs distributed Pareto of the second kind, that is* $X_i \sim \mathrm{Pa(II)}(\alpha, \theta_i)$*,* $\alpha \in \mathbb{R}_+$*,* $\theta_i \in \mathbb{R}_+$*,* $i = 1, 2$*, and assume that the RVs* X_1 *and* X_2 *are co-monotonic. Set* $g_1(\boldsymbol{x}) = x_1 + x_2$ *and* $g_2(\boldsymbol{x}) = \xi \circ g_1(\boldsymbol{x})$*, where* $\xi(y) = \max(0, y - d)$ *represents the risk reduction function due to the inclusion of a deductible* $d > 0$*. Furthermore, consider the same weight function as in Example 1, i.e.,* $w(x) = x^b$*,* $b \in \mathbb{R}_+$*, which is log-concave and thus violates the conditions on* w *in Theorem 4. Let* $S_g = g_1(\boldsymbol{X}) = X_1 + X_2$*, then since the loss RVs are co-monotonic, we have* $S_g \sim \mathrm{Pa(II)}(\alpha, \theta^*)$*, where* $\theta^* = \theta_1 + \theta_2$*. Finally, assume that the succeeding expectations are well-defined and finite, or equivalently* $\alpha > b + 1$*, and have*

$$\begin{aligned}
A_{i,1} &= \frac{\mathbb{E}\big[X_i \times w(S_g)\big]}{\mathbb{E}\big[w(S_g)\big]} \\
&= \frac{\mathbb{E}\big[X_i\,(X_1 + X_2)^b\big]}{\mathbb{E}\big[(X_1 + X_2)^b\big]} \\
&= \frac{\mathbb{E}\big[X_i^{1+b}\big]}{\mathbb{E}\big[X_i^b\big]} \\
&= \theta_i\,\frac{1+b}{\alpha - b - 1}, \qquad i = 1, 2.
\end{aligned}$$

Also, we have, for $i = 1, 2,$*,*

$$\begin{aligned}
A_{i,2} &= \frac{\mathbb{E}\big[X_i \times w \circ \xi(S_g)\big]}{\mathbb{E}\big[w \circ \xi(S_g)\big]} \\
&= \frac{\mathbb{E}\big[X_i\,(X_1 + X_2 - d)^b \,|\, X_1 + X_2 > d\big]}{\mathbb{E}\big[(X_1 + X_2 - d)^b \,|\, X_1 + X_2 > d/c\big]} \\
&= \frac{\theta_i}{\theta^*}\,\frac{\mathbb{E}\big[S_g\,(S_g - d)^b \,|\, S_g > d\big]}{\mathbb{E}\big[(S_g - d)^b \,|\, S_g > d\big]} \\
&= \frac{\theta_i}{\theta^*}\,\frac{\mathbb{E}\big[(S_g - d)^{b+1} \,|\, S_g > d\big] + d\,\mathbb{E}\big[(S_g - d)^b \,|\, S_g > d\big]}{\mathbb{E}\big[(S_g - d)^b \,|\, S_g > d\big]}.
\end{aligned}$$

Note that $S_g - d \,|\, S_g > d \sim \mathrm{Pa(II)}(\alpha, d + \theta^*)$*, and so we obtain*

$$A_{i,2} = \frac{\theta_i}{\theta^*}\Big[\big(d + \theta^*\big)\frac{1+b}{\alpha - b - 1} + d\Big] = A_{i,1} + \frac{\alpha\,d\,\theta_i}{(\alpha - b - 1)\,\theta^*} \geq A_{i,1}, \qquad i = 1, 2.$$

In conclusion, we have seen in this example that if the log-convexity assumption on the weight function w is violated, then the desired relationship $A_{i,1} \geq A_{i,2}$ reported in Theorem 4 is not guaranteed. Thereby, the order of the aggregation functions g_1 and g_2 is not sufficient to determine the order of the associated generalized weighted risk functionals. Moreover, if the weight function w is not log-convex, then (5) may fail to capture the risk reduction due to the introduction of policy modifications. This completes Example 2.

3 Orders based on different weight functions but the same aggregation function

We are now in a position to examine the role that the weight function plays in the determination of the order of risk functionals (4) and (5), given that they share the same aggregation function. We start with the study of the former generalized weighted risk functional, in which context our observations are summarized in the following theorem.

Theorem 5. *For an aggregation function $g : [0, +\infty)^n \to [0, +\infty)$, let*

$$\tilde{H}_j = H_{w_j}(S_g), \quad j = 1, 2$$

denote two generalized weighted risk functionals associated with the weight functions w_1 and w_2. Then the following relationships hold:

$$\text{If the ratio } y \mapsto \frac{w_1(y)}{w_2(y)},\ y \in \mathbb{R}_+,\ \text{is} \begin{Bmatrix} \text{non-decreasing} \\ \text{decreasing} \end{Bmatrix},\ \text{then } \tilde{H}_1 \begin{Bmatrix} \geq \\ < \end{Bmatrix} \tilde{H}_2.$$

Particularly, if $y \mapsto w_1(y)/w_2(y) \equiv c$, $y \in \mathbb{R}_+$, for some constant $c \in \mathbb{R}_+$, then $\tilde{H}_1 = \tilde{H}_2$.

Proof. The result follows from Theorem 4 of Patil and Rao (1978) and statement (4.3) of Furman and Zitikis (2008a). □

Interestingly, Theorem 5 shows that for a loss position $\boldsymbol{X} \in \mathcal{X}^n$ with a fixed aggregation function g, the monotonicity behaviour of the ratio of the two weight functions w_1 and w_2 may yield the order of the associated generalized weighted risk functionals. Furthermore, suppose that the g-aggregate RV S_g has a continuous CDF, then Table 2 summarizes the conditions under which the ratio $y \mapsto w_1(y)/w_2(y)$ is non-decreasing, and thus $H_{w_1}(S_g) \geq H_{w_2}(S_g)$ as per Theorem 5. Several observations pertaining to the conditions outlined in Table 2 are warranted and follow.

First, note that the net premium risk functional and the modified variance risk functional are special cases of the size-biased risk functional with $t = 0$ and $t = 1$, respectively (see, Table 1). Thereby, Table 2 can be immediately used to study the order of these two risk functionals.

Second, with the exception of the distortion functionals, the diagonal cells in Table 2 indicate that, for any two weight functions belonging to the same class, it is sufficient to use the value of the t parameter to determine the order of the associated weighted risk functionals.

Third, when comparisons are made across different classes of weight functions (and hence distinct weighed risk functionals), then the monotonicity behavior of the ratio $y \mapsto w_1(y)/w_2(y)$ may depend on the support and/or the probability distribution of the RV S_g, except for the comparison between the size-biased and Kamps' risk functionals. More specifically, for the comparison between the size-biased and Esscher's risk functionals, the ratio $y \mapsto w_1(y)/w_2(y)$ is non-decreasing (resp. decreasing) only when the g-aggregate RV S_g is bounded from above (resp. below) by t_1/t_2.

w_2 \ w_1	Size-biased $w_1(y) = y^{t_1}$	Esscher $w_1(y) = e^{t_1 y}$	Aumann-Shapley $w_1(y) = e^{t_1 F(y)}$	Kamps $w_1(y) = 1 - e^{-t_1 y}$	Distorted $w_1(y) = h_1^{'}\left(\overline{F}(y)\right)$
Size-biased $w_2(y) = y^{t_2}$	$t_1 \geq t_2$				
Esscher $w_2(y) = e^{t_2 y}$	$y \leq \frac{t_1}{t_2}$, $\forall y \in \mathbb{R}_+$	$t_1 \geq t_2$			
Aumann-Shapley $w_2(y) = e^{t_2 F(y)}$	$f(y)\, y \leq \frac{t_1}{t_2}$, $\forall y \in \mathbb{R}_+$	$f(y) \leq \frac{t_1}{t_2}$, $\forall y \in \mathbb{R}_+$	$t_1 \geq t_2$		
Kamps $w_2(y) = 1 - e^{-t_2 y}$	$t_1 \geq 1$	$y \geq \frac{1}{t_2} \log\left(\frac{t_2}{t_1} + 1\right)$, $\forall y \in \mathbb{R}_+$	$y \geq \frac{1}{t_2} \log\left(\frac{t_2}{t_1\, f(y)} + 1\right)$, $\forall y \in \mathbb{R}_+$	$t_2 \geq t_1$	
Distorted $w_2(y) = h_2^{'}\left(\overline{F}(y)\right)$	$h_2^{''}(y) \geq 0$, $\forall y \in (0,1)$	$h_2^{''}(y) \geq 0$, $\forall y \in (0,1)$	$h_2^{''}(y) \geq 0$, $\forall y \in (0,1)$	$h_2^{''}(y) \geq 0$ $\forall y \in (0,1)$	$h_1^{'}(y)h_2^{''}(y) \geq h_1^{''}(y)h_2^{'}(y)$ $\forall y \in (0,1)$

Table 2: Summary of the conditions such that the function $y \mapsto w_1(y)/w_2(y)$, $y \in \mathbb{R}_+$, is non-decreasing when the g-aggregate RV S_g has a continuous CDF. The ranges of the parameters are $t_1, t_2 \in \mathbb{R}_+$, and $f(y),\ y \in \mathbb{R}_+$ denotes the probability density function of S_g.

When it comes to the comparison between the size-biased and the Aumann-Shapley risk functionals, note that commonly used absolutely continuous distributions such as the ones with unbounded supports outlined in the distribution inventory of Klugman et al. (2012), have $y \mapsto f(y)\, y$ bounded for all $y \in \mathbb{R}_+$, where f denotes the respective PDF. Therefore, we can find sufficiently large t_1 and/or small t_2 such that the corresponding inequality condition is satisfied, and thus the ratio $w_1(y)/w_2(y)$ is non-decreasing. On a different note, it is also worth mentioning that $y \mapsto f(y)\, y$ is not always bounded from above. A counter example is the arcsine distribution, or more generally a Beta distribution with the second shape parameter being less than one, whose density is given by

$$f(y) = \frac{1}{\pi\sqrt{y(1-y)}}, \qquad y \in (0,1);$$

it is evident that $\lim_{y\uparrow 1} f(y)\, y = +\infty$. It is also possible that $y \mapsto f(y)\, y$ is bounded from below by a positive value (e.g., the right-shifted uniform distribution). In this case, we can find an appropriate pair of t_1 and t_2 such that $f(y)\, y \geq t_1/t_2$, thus $y \mapsto w_1(y)/w_2(y)$ is decreasing for all $y \in \mathbb{R}_+$. For such common absolutely continuous distributions as gamma, log-normal, Pareto, and Weibull, we have $\lim_{y\downarrow 0} f(y)\, y = 0$, thus it is impossible that $f(y)\, y \geq t_1/t_2$ for all $y \in \mathbb{R}_+$, and $y \mapsto w_1(y)/w_2(y)$ can not be decreasing.

Turning to the Esscher functional, its comparison with Kamps functional suggests that the support of the g-aggregate RV S_g must have a positive lower (resp. upper) bound ${t_2}^{-1} \log(t_2/t_1 + 1)$ such that $y \mapsto w_1(y)/w_2(y)$ is non-decreasing (resp. decreasing). To implement the comparison between the Esscher and Aumann-Shapley functionals, we require the density of S_g to be bounded from above or from below by a positive value. When the RV S_g has an unbounded support, then it is impossible that the PDF f has a positive lower bound, thus $y \mapsto w_1(y)/w_2(y)$ can not be non-decreasing.

Penultimately, let us consider the comparison between the Kamps and Aumann-Shapley functionals. The inequality for ensuring the non-decreasing behavior of w_1/w_2 depends on both the support of the RV S_g and the behaviour of the PDF f. If f is unbounded at a positive point, then $\log\big(t_2(t_1\, f(y))^{-1} + 1\big) \to 0$ as y approaches that point. So the corresponding inequality condition specified in Table 2 holds, and the non-decreasing property of w_1/w_2 can be established in a neighbourhood of the mentioned positive point. However, when the PDF $f(y)$ converges to zero as y approaches a finite point (e.g., the gamma distribution having the PDF $f(y) = \frac{1}{2}y^2e^{-y}$, $y \in \mathbb{R}_+$), then $\log\big(t_2(t_1\, f(y))^{-1} + 1\big) \to \infty$ as y approaches zero, and it is impossible that $y \geq \log\big(t_2(t_1\, f(y))^{-1} + 1\big)$ for all $y \in \mathbb{R}_+$; thus the function $y \mapsto w_1(y)/w_2(y)$ can not be non-decreasing.

Finally, the comparison between the distorted functional and all others is quite simple, as it can solely depend on the function h. The sufficient condition of positive second derivative, of h, ensures that $y \mapsto w_1(y)/w_2(y)$ is non-decreasing. As h is usually chosen non-decreasing, then $h^{'} \geq 0$ is automatically satisfied, then the only remaining restriction is that imposed on the second degree. If h is convex, meaning $h^{''} \geq 0$, then we get the desired monotonicity behaviour of the ratio. If we compare two distortion functionals then the behaviour of the product of the two derivatives determines the functional order.

In what follows, we proceed to studying the conditions for determining the order of generalized weighted risk functionals (5) subject to different weight functions but with a common aggregation function. To this end, we need the following additional notation

$$\tilde{\ell}_{i,j}(y) = \mathbb{E}[w_j(S_g)|X_i = y], \quad \text{where } S_g = g(\boldsymbol{X}),\ j = 1, 2,\ i \in \mathcal{N}, \text{ and } y \geq 0.$$

Theorem 6. *For an aggregation function g, let*

$$\tilde{A}_{i,j} = A_{w_j}(X_i, S_g),\ \textit{for } S_g = g(\boldsymbol{X}) \textit{ and } j = 1, 2,$$

denote the weighted risk functionals associated with the weight functions w_1 and w_2.

$$\textit{If the ratio } y \mapsto \frac{\tilde{\ell}_{i,1}(y)}{\tilde{\ell}_{i,2}(y)},\ y \in \mathbb{R}_+,\ \textit{is} \left\{\begin{matrix} \textit{non-decreasing} \\ \textit{decreasing} \end{matrix}\right\},\ \textit{then } \tilde{A}_{i,1} \left\{\begin{matrix} \geq \\ < \end{matrix}\right\} \tilde{A}_{i,2}.$$

In particular, if $\tilde{\ell}_{i,1}(y)/\tilde{\ell}_{i,2}(y) \equiv c$ for some constant $c \in \mathbb{R}_+$, then $\tilde{A}_{i,1} = \tilde{A}_{i,2}$.

Proof. The result follows from Proposition 3.1 of Furman and Zitikis (2008b). □

Interestingly, for co-montonic losses Theorem 6 simplifies significantly.

Theorem 7. *Let us consider* $\boldsymbol{X} = (X_1, \ldots, X_n)$ *with* $X_i \stackrel{d}{=} F_{X_i}^{-1}(U)$, *where* $U \sim \text{Uniform}(0,1)$ *and* F_{X_i} *is the CDF of the RV* X_i, $i \in \mathcal{N}$. *Fix a component-wise non-decreasing aggregation function* g, *then following relationships hold:*

$$\text{If the ratio } y \mapsto \frac{w_1(y)}{w_2(y)},\ y \in \mathbb{R}_+, \text{ is } \begin{Bmatrix} \text{non-decreasing} \\ \text{decreasing} \end{Bmatrix}, \text{ then } \tilde{A}_{i,1} \begin{Bmatrix} \geq \\ < \end{Bmatrix} \tilde{A}_{i,2}.$$

In particular, if $w_1(y)/w_2(y) \equiv c$ *for some constant* $c \in \mathbb{R}_+$, *then* $\tilde{A}_{i,1} = \tilde{A}_{i,2}$.

Proof. The proof of Theorem 7 is relegated to Appendix Aiv. □

Theorem 7 seems to suggest that the monotonicity behavior of ratios of the weight functions may be a decisive factor as to the orders of risk functionals (5), as it is in the context of risk functionals (4). The next example shows that it is not the case, if the co-monotonicity assumption on the losses of interest is lifted.

Example 3. *Consider the loss RV* $\boldsymbol{X} = (X_1, X_2)$, *whose probabilistic behavior is governed by a two-component mixture of gamma distributions with the joint PDF (Chen et al., 2021)*

$$f_{X_1,X_2}(x_1,x_2) = p \prod_{i=1}^{2} \frac{x_i^{\alpha_{i1}-1}\theta_i^{\alpha_{i1}}}{\Gamma(\alpha_{i1})} e^{-\theta_i x_i} + (1-p) \prod_{i=1}^{2} \frac{x_i^{\alpha_{i2}-1}\theta_i^{\alpha_{i2}}}{\Gamma(\alpha_{i2})} e^{-\theta_i x_i}, \qquad x_1, x_2 \in \mathbb{R}_+,\ p \in (0,1).$$

In this example, we set the aggregation function $g(\boldsymbol{x}) = x_1 + x_2$, *and consider two weight functions* $w_j(y) = y^{n_j}$, *where* $n_j \in \mathbb{N}$, $j = 1,\ 2$, *with* $n_1 \geq n_2$. *Clearly, the ratio* $y \mapsto w_1(y)/w_2(y)$ *is non-decreasing. Next, let us fix* $p = 0.5$, $\alpha_{11} = 2$, $\alpha_{12} = 1$, $\alpha_{21} = \alpha > 0$, *and* $\alpha_{22} = 8$. *Also, let* $n_1 = 2$ *and* $n_2 = 1$. *Figure 1 depicts the Pearson correlation of the pair of losses,* (X_1, X_2), *and the corresponding weighted risk functionals (5) as functions of* α, *which are computed based on Corollary 3 and Proposition 2 of (Chen et al., 2021). As observed, the order* $\tilde{A}_{1,1} < \tilde{A}_{1,2}$ *holds for smaller* $\alpha \in \mathbb{R}_+$. *Hence, the non-decreasing property of* w_1/w_2 *is not sufficient to yield the desired order* $\tilde{A}_{1,1} \geq \tilde{A}_{1,2}$ *as per Theorem 7 after the co-monotonicity assumption is lifted. Nevertheless, as the value of the parameter* α *increases, the loss RVs* X_1 *and* X_2 *become more positively correlated, as demonstrated by the non-decreasing pattern of the Pearson correlation, and the order between* $\tilde{A}_{1,1}$ *and* $\tilde{A}_{1,2}$ *tends to coincide with the one suggested by Theorem 7 for co-monotonic losses.*

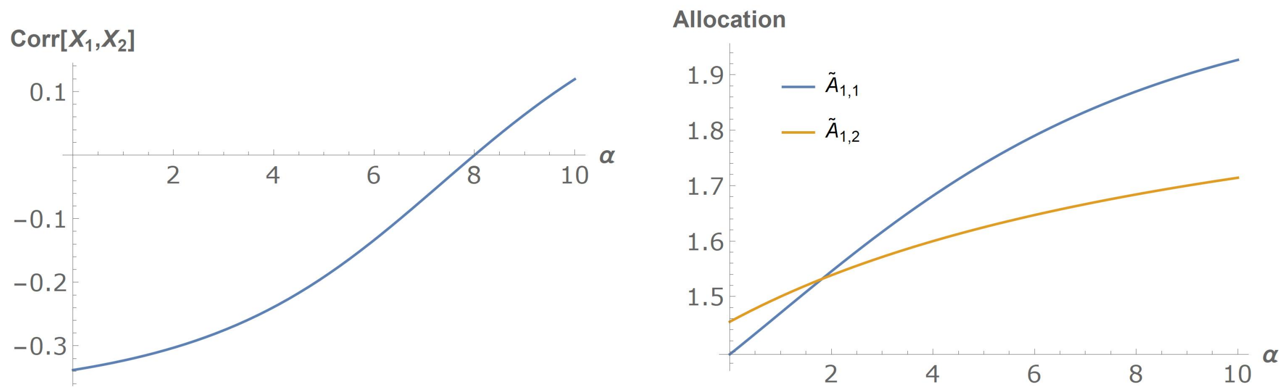

Figure 1: Plots of the Pearson correlation of the pair of loss RVs (X_1, X_2) and weighted risk functionals (5) as functions of $\alpha \in (0, 10)$.

In summary, the study in this section suggests that while the monotonicity behavior of the ratio of two weight functions, $y \mapsto w_1(y)/w_2(y)$, may play a decisive role in the determination of the orders of risk functionals (4), this is not generally the case in the context of risk functionals (5).

4 Afterthoughts and related results

The results we have established thus far suggest that the choices of the aggregation function, g, and the weight function, w, have rather complex interactive effects on the orderings of the generalized weighted risk functionals. Therefore, the precise values of functionals (4) and (5) may be frequently needed to be computed in applications. We note in this respect that the determination of the value of the latter generalized weighted risk functional is in general significantly more involved than the determination of the value of the former generalized weighted risk functional. That being said, there exist loss RVs $\boldsymbol{X} \in \mathcal{X}^n$, for which the two mentioned exercises turn out to be of the same complexity. Such special loss RVs are discussed in this section.

At the outset, we observe the following identity, for $i \in \mathcal{N}$,

$$A_w(X_i, S_g) = \mathbb{E}[X_i] + \frac{\mathrm{Cov}(X_i, w(S_g))}{\mathrm{Cov}(S_g, w(S_g))} \left(H_w(S_g) - \mathbb{E}[S_g]\right), \tag{7}$$

where we routinely assume that all the involved quantities are well-defined and finite. Equation (7) is reminiscent of the renowned Capital Asset Pricing Model. For this reason, Furman and Zitikis (2010) call a special case of (7), in which $S_g = X_1 + \cdots + X_n$, the Weighted Insurance Pricing Model (WIPM).

Further, for some non-negative α_i and $\beta_i,\ i \in \mathcal{N}$, let us define the following collection of loss RVs

$$\mathcal{T} = \left\{ \boldsymbol{X} \in \mathcal{X}^n : \, A_w(X_i, S_g) = \alpha_i + \beta_i \times H_w(S_g) \text{ for any choice of } w \text{ and } i \in \mathcal{N} \right\};$$

here α_i and β_i do not depend on the weight function w. Also, for $\alpha_i \equiv 0$ and $\beta_i \equiv 0$, the particular cases of the set $\mathcal{T}$ are in what follows denoted by $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. The loss RVs that belong to the set $\mathcal{T}$ are such that the complexity of computing generalized weighted risk functionals (4) and (5) is the same. Moreover, the loss RVs that comprise the set $\mathcal{T}_1$ imply $A_w(X_i, w(S_g))/H_w(S_g) = \beta_i,\ i \in \mathcal{N}$, and we thus refer to this phenomenon as 'proportional triviality'. Finally, the set $\mathcal{T}_2$ contains those loss RVs $\boldsymbol{X} \in \mathcal{X}^n$, for which $A_w(X_i, w(S_g)) = \alpha_i,\ i \in \mathcal{N}$, a phenomenon that is naturally refered to as 'absolute triviality'.

The characterization of the distributions of the loss RVs in the collection of losses $\mathcal{T}$ for the case of the simple sum aggregation function is studied in Furman and Zitikis (2008b) and in a more recent Mohammed et al. (2021). As therein, the regression function $y \mapsto \mathbb{E}[X_i |\ S_g = y],\ y \in \mathbb{R}_+$ plays in our deliberations an important role, as readily follows from the following auxiliary statement.

Assume that

$$\mathbb{P}\big(S_g \in [0, \epsilon)\big) > 0 \text{ for any } \epsilon > 0, \tag{8}$$

and hence $\mathbb{E}[X_i |\ S_g \in [0, \epsilon)]$ is well-defined.

Lemma 8. *For a fixed aggregation function g and the aggregate loss RV $S_g = g(\boldsymbol{X})$, we have $\boldsymbol{X} \in \mathcal{T}$ if and only if*

$$\mathbb{E}[X_i \,|\, S_g] \equiv \alpha_i + \beta_i \, S_g, \quad i \in \mathcal{N}. \tag{9}$$

Proof. The proof is kindred to the proof of Theorem 3.1 in Furman et al. (2018) and is thus omitted. □

In order to state and prove the main result of this section, we need the following notion of the ‘multivariate size-biased transform’.

Definition 1. *Let* $\boldsymbol{X} \in \mathcal{X}^n$ *be a loss RV with positive univariate coordinates* X_i, $i \in \mathcal{N}$ *having finite means. Then the multivariate coordinate-wise size-biased counterpart of* $\boldsymbol{X}$*, denoted by* $\boldsymbol{X}^{(i)}$*, is*

$$\mathbb{P}\big(\boldsymbol{X}^{(i)} \in \mathrm{d}\boldsymbol{x}\big) = \frac{x_i}{\mathbb{E}[X_i]}\,\mathbb{P}\big(\boldsymbol{X} \in \mathrm{d}\boldsymbol{x}\big) \text{ for all } \boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}_+^n. \tag{10}$$

Here the RVs $\boldsymbol{X}$ *and* $\boldsymbol{X}^{(i)}$ *are independent.*

Remark 3. *When* $n = 1$*, then the multivariate size-biased transform considered in Definition 1 reduces to the classical notion of univariate size-biased transforms (Patil and Rao, 1978). Namely, the size-biased counterpart of the RV* $X \in \mathcal{X}$ *with* $\mathbb{E}[X] < +\infty$*, denoted by* X^**, is such that*

$$\mathbb{P}\big(X^* \in \mathrm{d}x\big) = \frac{x}{\mathbb{E}[X]}\,\mathbb{P}\big(X \in \mathrm{d}x\big) \text{ for all } x \in \mathbb{R}_+.$$

The RVs X *and* X^* *are independent.*

We are now ready to formulate and prove the main statement of this section. (When the aggregation function g is the simple sum aggregation function, i.e., $g(\boldsymbol{x}) = x_1 + \cdots + x_n$, then the conclusions of Theorem 9 are analogous to the ones of Theorem 1 in Mohammed et al. (2021).)

Assume

$$g(x_1, \ldots, x_n) \downarrow 0 \quad \text{only when } \max(x_1, \ldots, x_n) \downarrow 0, \tag{11}$$

which ensures that no positive losses of a portfolio of losses are neglected when the aggregation function is zero.

Theorem 9. *Consider a loss RV* $\boldsymbol{X} \in \mathcal{X}^n$ *that has positive univariate coordinates* X_i, $i \in \mathcal{N}$ *with finite means. Fix an aggregation function* g *and suppose that the conditions in (8) and (11) hold implying that* $\alpha_i \equiv 0$*. Then* $\boldsymbol{X} \in \mathcal{T}_1$ *if and only if*

$$S_g{}^{(1)} \stackrel{d}{=} S_g{}^{(2)} \stackrel{d}{=} \cdots \stackrel{d}{=} S_g{}^{(n)}, \tag{12}$$

where ${S_g}^{(i)} := g\big(\boldsymbol{X}^{(i)}\big)$, $i \in \mathcal{N}$. *Further, we have* ${S_g}^{(1)} \stackrel{d}{=} {S_g}^*$, *where* ${S_g}^*$ *is the size-biased counterpart of* S_g. *Furthermore, it must hold that* $\beta_i = \mathbb{E}[X_i] \,/\, \mathbb{E}[S_g]$, $i \in \mathcal{N}$.

Proof. The proof of Theorem 9 is relegated to Appendix Av. □

It is evident that if the loss RV $\boldsymbol{X} \in \mathcal{X}^n$ is exchangeable, and the aggregation function is symmetric, then distributional equalities (12) hold. If this is the case, the $A_w(X_i, S_g) = 1/n$ for all $i \in \mathcal{N}$. Thereby, (12) implies certain symmetric structure inherent in both the distribution of the loss RV $\boldsymbol{X}$ and in the choice of the aggregation function. When the notion of multivariate size-biased transform as per Definition 1 is interpreted in terms of loading for, e.g., model risk, then equalities (12) signify that the choice of the loading direction does not impact the end-distribution of the aggregate loss.

One may now wonder under which conditions the other (extremal) case, in which $\alpha_i \in \mathbb{R}_+$ and $\beta_i \equiv 0$, holds. Namely, we are interested in studying the collection of losses $\mathcal{T}_2$ or the case of absolute – rather than proportional as in Theorem 9 – triviality.

Theorem 10. *Consider a loss RV* $\boldsymbol{X} \in \mathcal{X}^n$. *Suppose that none of the coordinates of the RV* $\boldsymbol{X}$ *is degenerate (i.e.,* $\mathbb{P}(X_i = \text{const}) < 1$, $i \in \mathcal{N}$*). Then it holds that* $\boldsymbol{X} \in \mathcal{T}_2$ *if and only if*

$$\mathbb{P}(S_g = c) = 1 \text{ with } S_g = g(\boldsymbol{X}) \text{ and a positive constant } c. \tag{13}$$

Proof. The proof of Theorem 10 to Appendix Avi. □

While Theorem 9 shows that proportional triviality allows for certain degree of richness of the corresponding class of distributions of the loss RVs $\boldsymbol{X} \in \mathcal{T}_1$, the case of absolute triviality in Theorem 10 is notably more restrictive with the condition $g(\boldsymbol{X}) = c$ almost surely for a constant $c \in \mathbb{R}_+$ for the loss RVs $\boldsymbol{X} \in \mathcal{T}_2$. For both trivialities, nonetheless, the underlying central condition is the regression function $y \mapsto \mathbb{E}[X_i | S_g = y]$ being either linear or constant in $y \in \mathbb{R}_+$. The regression condition draws similarities with, e.g., Guan et al. (2021), where an axiomatic formulation, in particular the axiom of shrinking independence, is used to reach absolute triviality.

Going back to Theorem 10, we note that the loss RVs that belong to the collection of losses $\mathcal{T}_2$ are referred to as g-joint mix RVs and are those RVs that have their CDFs supported on non-increasing sets (Bignozzi and Puccetti, 2015).

5 Conclusions

In this paper, we have introduced a generalized version of the nowadays well-studied class of weighted risk functionals. The generalization herein arises from the recognition that insurers' aggregate financial loss position must not be a result of a simple sum of losses, but rather can be specified by a general aggregation function. Loss aggregation in internal models is one example that elucidates the importance of general aggregation mechanisms, whereas coverage modifications are another example.

Naturally, generalizations bode complexity, and we admit that the proposed generalized weighted risk functionals are not easy to deal with. Nevertheless, we have identified a number of sufficient conditions, under which, the generalized weighted risk functionals can be ordered with respect to the choices of the aggregation functions and the weight functions, thus providing valuable insights to risk professionals. Moreover, we have characterized those loss RVs, for which the evaluation of the generalized weighted risk functionals - irrespective of the aggregation and weight functions - is of surprising simplicity.

Appendix A Proofs

i Proof of Theorem 1

Proof. We only prove the first case in which $h(y) \geq y$ and the function $y \mapsto y/\tilde{h}(y)$ is non-decreasing. The other case holds based on the same argument. Let us write

$$H_1 = \frac{\mathbb{E}\big[S_{g_1} \times w(S_{g_1})\big]}{\mathbb{E}\big[w(S_{g_1})\big]} = \frac{\mathbb{E}_{\tilde{h}\circ w(S_{g_1})}\big[S_{g_1} \times w(S_{g_1}) \times \big(\tilde{h}\circ w(S_{g_1})\big)^{-1}\big]}{\mathbb{E}_{\tilde{h}\circ w(S_{g_1})}\big[w(S_{g_1}) \times \big(\tilde{h}\circ w(S_{g_1})\big)^{-1}\big]}.$$

Since $y \mapsto y/\tilde{h}(y)$ is non-decreasing, then using Chebyshev's sum inequality we have

$$\mathbb{E}_{\tilde{h}\circ w(S_{g_1})}\big[S_{g_1} \times w(S_{g_1}) \times \big(\tilde{h}\circ w(S_{g_1})\big)^{-1}\big] \geq \mathbb{E}_{\tilde{h}\circ w(S_{g_1})}[S_{g_1}]$$

$$\times\, \mathbb{E}_{\tilde{h}\circ w(S_{g_1})}\big[w(S_{g_1}) \times \big(\tilde{h}\circ w(S_{g_1})\big)^{-1}\big].$$

Thereby, it holds that

$$\begin{aligned} H_1 \geq \mathbb{E}_{\tilde{h}\circ w(S_{g_1})}[S_{g_1}] &= \frac{\mathbb{E}\big[S_{g_1} \times w(S_{g_2})\big]}{\mathbb{E}\big[w(S_{g_2})\big]} = \frac{\mathbb{E}\big[h(S_{g_2}) \times w(S_{g_2})\big]}{\mathbb{E}\big[w(S_{g_2})\big]} \\ &\geq \frac{\mathbb{E}\big[S_{g_2} \times w(S_{g_2})\big]}{\mathbb{E}\big[w(S_{g_2})\big]} = H_2, \end{aligned}$$

where the second inequality holds because of the condition $h(y) \geq y$, $y \in \mathbb{R}_+$. The proof is now completed. □

ii Proof of Theorem 2

Proof. We only prove the first case in which $\xi(y) \leq y$ with $\xi'(y) \leq 1$ for all $y \in \mathbb{R}_+$. A repeated application of the same argument yields the desired conclusion for the second case.

Recall that $S_{g_j} = g_j(\boldsymbol{X})$ for $j = 1, 2$. Since $\xi \circ g_1(\boldsymbol{y}) = g_2(\boldsymbol{y})$, we can write

$$\frac{w(y)}{\mathbb{E}\big[w(S_{g_2}) \,|\, S_{g_1} = y\big]} = \frac{w(y)}{w \circ \xi(y)},$$

which has the same monotonicity behavior as $y \mapsto \log\big(w(y)\big) - \log\big(w \circ \xi(y)\big)$, $y \in \mathbb{R}_+$. Consider

$$\frac{\mathrm{d}}{\mathrm{d}y}\Big[\log\big(w(y)\big) - \log\big(w \circ \xi(y)\big)\Big] = \frac{\mathrm{d}}{\mathrm{d}t}\log\big(w(t)\big)\big|_{t=y} - \frac{\mathrm{d}}{\mathrm{d}t}\log\big(w(t)\big)\big|_{t=\xi(y)}\,\xi'(y).$$

By assumption, we have $\xi'(y) \in (0, 1]$ for $y \in \mathbb{R}_+$. Moreover, since $\xi(y) \leq y$ and $t \mapsto w(t)$ is log-convex, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\big(w(t)\big)\big|_{t=y} \geq \frac{\mathrm{d}}{\mathrm{d}t}\log\big(w(t)\big)\big|_{t=\xi(y)}.$$

Together with the assumption that $y \mapsto w(y)$ is non-decreasing on $y \in \mathbb{R}_+$, we conclude

$$\frac{\mathrm{d}}{\mathrm{d}y}\Big[\log\big(w(y)\big) - \log\big(w \circ \xi(y)\big)\Big] \geq 0,$$

thus the function $y \mapsto w(y)/\mathbb{E}[w(S_{g_2}) \mid S_{g_1} = y]$ is non-decreasing on $y \in \mathbb{R}_+$.

Further, note that $\xi(y) \le y$ implies $h(y) \ge y$ for all $y \in \mathbb{R}_+$ based on Remark 1. According to Theorem 1 and Remark 2, we obtain $H_1 \ge H_2$. The proof is now completed. □

iii Proof of Theorem 4

Proof. Let $S_g = g_1(\boldsymbol{X})$, then the following string of equations holds, for $x \in \mathbb{R}_+$,

$$\begin{aligned}\frac{\ell_{i,1}(x)}{\ell_{i,2}(x)} &= \frac{\mathbb{E}\big[w(S_g)|X_i = x\big]}{\mathbb{E}\big[w \circ \xi(S_g)|X_i = x\big]} \\ &= \frac{\mathbb{E}\big[w(S_g)|U = F_{X_i}(x)\big]}{\mathbb{E}\big[w \circ \xi(S_g)|U = F_{X_i}(x)\big]} \\ &= \frac{w \circ \tilde{g}(x)}{w \circ \xi \circ \tilde{g}(x)},\end{aligned} \tag{14}$$

where $\tilde{g}(x) = g_1(\boldsymbol{x})$ with $\boldsymbol{x} = \big(F_{X_j}^{-1}(F_{X_i}(x))\big)_{j \in \mathcal{N}}$.

Since the aggregation function is element-wise non-decreasing, the monotonicity behaviour of the ratio in (14) is same as that of $y \mapsto w(y)/w \circ \xi(y)$. Evoking the argument used in the proof of Theorem 2, we conclude that $x \mapsto \ell_{i,1}(x)/\ell_{i,2}(x)$ is non-decreasing if $\xi(y) \le y$ and $\xi'(y) \le 1$, and the function is decreasing if $\xi(y) > y$ and $\xi'(y) > 1$. Applying Theorem 3 yields the desired result and thus completes the proof. □

iv Proof of Theorem 7

Proof. Fix an aggregation function g. Then let us write

$$\begin{aligned}\frac{\tilde{\ell}_{i,1}(x)}{\tilde{\ell}_{i,2}(x)} &= \frac{\mathbb{E}\big[w_1(S_g)|X_i = x\big]}{\mathbb{E}\big[w_2(S_g)|X_i = x\big]} \\ &= \frac{\mathbb{E}\big[w_1 \circ g(\boldsymbol{X})|U = F_{X_i}(x)\big]}{\mathbb{E}\big[w_2 \circ g(\boldsymbol{X})|U = F_{X_i}(x)\big]} \\ &= \frac{w_1 \circ \tilde{g}(x)}{w_2 \circ \tilde{g}(x)},\end{aligned} \tag{15}$$

where $\tilde{g}(x) = g(\boldsymbol{x})$ with $\boldsymbol{x} = \big(F_{X_k}^{-1}(F_{X_i}(x))\big)_{k \in \mathcal{N}}$.

Since the aggregation function g is component-wise non-decreasing, the function $x \mapsto \tilde{g}(x)$ is non-decreasing. We can conclude that the ratio $x \mapsto \tilde{\ell}_{i,1}(x)/\tilde{\ell}_{i,2}(x)$ has the same monotonicity behavior as the ratio $y \mapsto w_1(y)/w_2(y)$. An application of Theorem 6 yields the desired result, which completes the proof. □

v Proof of Theorem 9

Proof. Let us begin with the sufficiency of the statement. First note that conditions (8) and (11) together imply $\mathbb{E}[X_i|S_g = y] \downarrow 0$ when $y \downarrow 0$, $i \in \mathcal{N}$. Then evoking Lemma 8 yields that if $\boldsymbol{X} \in \mathcal{T}$, then α_i in (9) must be zero, i.e. $\boldsymbol{X} \in \mathcal{T}_1$, and thus

$$\mathbb{E}[X_i|S_g] = \beta_i\, S_g, \quad i \in \mathcal{N}.$$

Taking expectations throughout thus implies $\beta_i = \mathbb{E}[X_i] \,/\, \mathbb{E}[S_g]$.

The Laplace transform of the RV ${S_g}^{(i)}$ can be computed via, for $t \in \mathbb{R}_+$,

$$\begin{aligned} \frac{\mathbb{E}\left[X_i\, e^{-tS_g}\right]}{\mathbb{E}[X_i]} &= \frac{\mathbb{E}\left[\mathbb{E}[X_i|S_g]\, e^{-tS_g}\right]}{\mathbb{E}[X_i]} \\ &= \frac{\beta_i}{\mathbb{E}[X_i]}\, \mathbb{E}\left[S_g\, e^{-tS_g}\right] \\ &= \frac{\mathbb{E}\left[S_g e^{-tS_g}\right]}{\mathbb{E}[S_g]}, \end{aligned}$$

which readily implies the desirable equalities in distribution, that is ${S_g}^{(1)} \stackrel{d}{=} {S_g}^{(2)} \stackrel{d}{=} \cdots \stackrel{d}{=} {S_g}^{(n)} \stackrel{d}{=} {S_g}^{*}$.

To prove the necessity direction, note that for any $t \in \mathbb{R}_+$, we have

$$\frac{\mathbb{E}\left[\mathbb{E}[X_i|S_g]\, e^{-tS_g}\right]}{\mathbb{E}[X_i]} - \frac{\mathbb{E}\left[S_g e^{-tS_g}\right]}{\mathbb{E}[S_g]} = 0$$

is equivalent to

$$\mathbb{E}\left[\left(\mathbb{E}[X_i|S_g] - \frac{\mathbb{E}[X_i]}{\mathbb{E}[S_g]} S_g\right) e^{-tS_g}\right] = 0, \tag{16}$$

which in turn yields

$$\mathbb{E}[X_i|S_g] = \frac{\mathbb{E}[X_i]}{\mathbb{E}[S_g]}\, S_g,$$

and hence $\boldsymbol{X} \in \mathcal{T}_1$. This completes the proof. $\square$

vi Proof of Theorem 10

Proof. Based on Lemma 8 with $\beta_i = 0$, $i \in \mathcal{N}$, we obtain that $\boldsymbol{X} \in \mathcal{T}_2$ is equivalent to $\mathbb{E}[X_i|S_g] = \alpha_i$ and $\alpha_i = \mathbb{E}[X_i]$.

In order to prove the sufficiency direction, consider the Laplace transform of the RV ${S_g}^{(i)}$, that is

$$\begin{aligned}\mathbb{E}\left[e^{-t{S_g}^{(i)}}\right] &= \frac{\mathbb{E}\left[X_i\, e^{-tS_g}\right]}{\mathbb{E}[X_i]} = \frac{\mathbb{E}\left[\mathbb{E}[X_i|S_g]\, e^{-tS_g}\right]}{\mathbb{E}[X_i]} \\ &= \frac{\alpha_i}{\mathbb{E}[X_i]}\, \mathbb{E}\left[e^{-tS_g}\right] = \mathbb{E}\left[e^{-tS_g}\right], \quad t \in \mathbb{R}_+. \end{aligned} \tag{17}$$

Since the RVs X_i are non-degenerate, (17) implies $\mathbb{P}(S_g = c) = 1$ for some constant $c \in \mathbb{R}_+$. The implication holds since otherwise ${S_g}^{(i)} \stackrel{d}{=} S_g$, which leads to a contradiction.

To prove the necessity direction, note that for any $t \in \mathbb{R}_+$, we have

$$\frac{\mathbb{E}\left[\mathbb{E}[X_i|S_g]\, e^{-tS_g}\right]}{\mathbb{E}[X_i]} - \mathbb{E}\left[e^{-tS_g}\right] = 0,$$

which is equivalent to

$$\mathbb{E}\Big[\big(\mathbb{E}[X_i|S_g] - \mathbb{E}[X_i]\big)\, e^{-tS_g}\Big] = 0.$$

This implies

$$\mathbb{E}[X_i|S_g] = \mathbb{E}(X_i),$$

or equivalently, $\boldsymbol{X} \in \mathcal{T}_2$. This completes the proof of the theorem. $\square$

References

Bignozzi, V. and Puccetti, G. (2015). Studying mixability with supermodular aggregating functions. *Statistics & Probability Letters*, 100:48–55.

Castano-Martinez, A., Lopez-Blazquez, F., Pigueiras, G., and Sordo, M. A. (2020). A method for constructing and interpreting some weighted premium principles. *ASTIN Bulletin: The Journal of the IAA*, 50(3):1037–1064.

Chen, Y., Song, Q., and Su, J. (2021). On a class of multivariate mixtures of gamma distributions: Actuarial applications and estimation via stochastic gradient methods. *Variance*.

Choo, W. and de Jong, P. (2009). Loss reserving using loss aversion functions. *Insurance: Mathematics and Economics*, 45(2):271–277.

Dhaene, J., Tsanakas, A., Valdez, E. A., and Vanduffel, S. (2012). Optimal capital allocation principles. *Journal of Risk and Insurance*, 79(1):1–28.

Furman, E., Kuznetsov, A., and Zitikis, R. (2018). Weighted risk capital allocations in the presence of systematic risk. *Insurance: Mathematics and Economics*, 79:75–81.

Furman, E. and Zitikis, R. (2007). An actuarial premium pricing model for nonnormal insurance and financial risks in incomplete markets, zinoviy landsman and michael sherris, january 2007. *North American Actuarial Journal*, 11(3):174–176.

Furman, E. and Zitikis, R. (2008a). Weighted premium calculation principles. *Insurance: Mathematics and Economics*, 42(1):459–465.

Furman, E. and Zitikis, R. (2008b). Weighted risk capital allocations. *Insurance: Mathematics and Economics*, 43(2):263–269.

Furman, E. and Zitikis, R. (2009). Weighted pricing functionals with applications to insurance: An overview. *North American Actuarial Journal*, 13(4):483–496.

Furman, E. and Zitikis, R. (2010). General Stein-type covariance decompositions with applications to insurance and finance. *ASTIN Bulletin*, 40(1):369–375.

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009). *Aggregation Functions*. Cambridge University Press.

Guan, Y., Tsanakas, A., and Wang, R. (2021). An impossibility theorem on capital allocation. *SSRN*.

Guo, Q., Bauer, D., and Zanjani, G. (2018). Capital allocation techniques: Review and comparison. Technical report, Variance.

Jaworski, P., Durante, F., Härdle, W. K., and Rychlik, T., editors (2010). *Copula Theory and Its Applications*, volume 198 of *Lecture Notes in Statistics*. Springer, Heidelberg.

Kaluszka, M. and Krzeszowiec, M. (2012). Pricing insurance contracts under cumulative prospect theory. *Insurance: Mathematics and Economics*, 50(1):159–166.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss Models: From Data to Decisions (4th ed.)*. Wiley, Hoboken.

Lehmann, E. L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics*, 37(5):1137–1153.

Millossovich, P., Tsanakas, A., and Wang, R. (2021). A theory of multivariate stress testing. *SSRN Electronic Journal*.

Mohammed, N., Furman, E., and Su, J. (2021). Can a regulatory risk measure induce profit-maximizing risk capital allocations? the case of conditional tail expectation. *Insurance: Mathematics and Economics*, 101:425–436.

Patil, G. P. and Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, pages 179–189.

Porth, L., Zhu, W., and Tan, K. S. (2014). A credibility-based erlang mixture model for pricing crop reinsurance. *Agricultural Finance Review*, 74(2):162–187.

Richards, D. and Uhler, C. (2019). Loading monotonicity of weighted premiums, and total positivity properties of weight functions. *Journal of Mathematical Analysis and Applications*, 475(1):532–553.

Sendov, H. S., Wang, Y., and Zitikis, R. (2011). Log-supermodularity of weight functions, ordering weighted losses, and the loading monotonicity of weighted premiums. *Insurance: Mathematics and Economics*, 48(2):257–264.

Zhu, W., Tan, K. S., and Porth, L. (2019). Agricultural insurance ratemaking: Development of a new premium principle. *North American Actuarial Journal*, 23(4):512–534.