Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks

Edward Furman

Department of Mathematics and Statistics, York University, Toronto, Ontario M3J 1P3, Canada. E-mail: efurman@mathstat.yorku.ca

Ruodu Wang[1]

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, N2L 5A7, Canada. E-mail: wang@uwaterloo.ca

Ričardas Zitikis

Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario N6A 5B7, Canada. E-mail: zitikis@stats.uwo.ca

Abstract. We introduce and explore Gini-type measures of risk and variability, and develop the corresponding economic capital allocation rules. The new measures are coherent, additive for co-monotonic risks, convenient computationally, and require only finiteness of the mean. To elucidate our theoretical considerations, we derive closed-form expressions for several parametric families of distributions that are of interest in insurance and finance, and further apply our findings to a risk portfolio of a bancassurance company.

JEL classification: C6, G10

Key words and phrases: risk measure, variability measure, Gini shortfall, Gini capital allocation, Choquet integral.

1 Introduction

Measuring risk is of pivotal importance in insurance and general finance. Not surprisingly, therefore, a large number of risk measures have been proposed and explored in the literature, which is abundant. The 2007–2009 financial crisis revived the interest in the notion of prudence in the regulatory frameworks for insurance and banking sectors (e.g., Cruz (2009), Sandström (2010), Cannata and Quagliariello (2011), Embrechts et al. (2014), and the references therein). As a result, a prominent trend associated with tail-based

[1]Corresponding author. Phone number: (+1)519-888-4567 ext. 31569

risk measures has emerged, with the value-at-risk (VaR) and the expected shortfall (ES) being arguably the most popular nowadays tail-based risk measures.

The VaR is a quantile, that is, given a prudence level $p \in (0,1)$ and a risk random variable (rv) X, whose cumulative distribution function (cdf) we denote by F_X, the value-at-risk $\mathrm{VaR}_p(X)$ is the p-th quantile of F_X given by

$$\mathrm{VaR}_p(X) = \inf\{x \in \mathbb{R} : F_X(x) \geq p\}. \tag{1.1}$$

The ES is the average of VaR over large prudence levels, that is,

$$\mathrm{ES}_p(X) = \frac{1}{1-p} \int_p^1 \mathrm{VaR}_q(X) \mathrm{d}q. \tag{1.2}$$

When the cdf F_X is continuous, then the ES risk measure coincides with the tail conditional expectation (TCE) risk measure, which is given by

$$\mathrm{TCE}_p(X) = \mathbb{E}[X \mid X > x_p], \tag{1.3}$$

where $\mathbb{E}$ denotes the expectation operator, and $x_p = \mathrm{VaR}_p(X)$; from here to the end of the section we assume $\mathbb{P}(X > x_p) > 0$ so that (1.3) is properly defined. Throughout the paper we interchangeably use the notation $\mathrm{VaR}_p(X)$, x_p and $F_X^{-1}(p)$ for the p-th quantile, depending on the tradition or notational simplicity.

Due to their nature, the above risk measures do not capture the *variability* of the risk rv X beyond the quantile x_p, yet the notion of variability in risk assessment has been prominent since at least 1952 when Harry Markowitz published his celebrated "Portfolio Selection" (Markowitz (1991)). To incorporate variability in tail risk analysis, Furman and Landsman (2006a) suggested the tail-standard-deviation (TSD) risk measure

$$\mathrm{TSD}_p^\lambda(X) = \mathrm{TCE}_p(X) + \lambda\, \mathrm{SD}_p(X), \tag{1.4}$$

where $p \in (0,1)$ is the prudence level, $\lambda \geq 0$ is the loading parameter, and the (tail) standard-deviation measure $\mathrm{SD}_p(X)$ is given by the equation

$$\mathrm{SD}_p(X) = \sqrt{\mathbb{E}[(X - \mathrm{TCE}_p(X))^2 \mid X > x_p]}. \tag{1.5}$$

When TCE is replaced by ES in equations (1.4) and (1.5), we call the resulting risk measure the standard-deviation shortfall (SDS) and denote it by $\mathrm{SDS}_p^\lambda(X)$. Obviously, the TSD and SDS risk measures may only be different if the cdf F_X is discontinuous. Unfortunately, the two risk measures TSD and SDS lack some crucial properties, such as:

- their definitions require finite second moments of the underlying risk rv's, and thus seriously impede the practical applicability. Indeed, plenty of evidence has come to light suggesting that the risks in insurance and finance often have infinite variance and finite mean (see e.g. Seal (1980) and Rachev (2003), respectively);

- they are not monotone, which contradicts the natural intuition behind the economic capital regulation, i.e., the smaller the risk, the less capital is required to make the risky position acceptable;

- they are not additive for co-monotonic risks. Additivity for co-monotonic risks means no diversification benefits rewarded to the aggregation of co-monotonic risks (e.g. Emmer et al. (2015)), and this property is satisfied by the practical risk measures VaR and ES;

- the TSD and SDS risk measures are undefined on some discrete risks violating the requirement $\mathbb{P}(X > x_p) > 0$.

In the present paper, therefore, we set out to develop an alternative way for measuring variability so that the resulting risk measures would be well defined for risks with infinite variances, monotone, and co-monotonically additive. These requirements naturally lead us to Gini-type measures of risk and variability, that we introduce and discuss below, thus providing an informative complement to the classical risk assessment based on the ubiquitous value-at-risk and expected shortfall risk measures.

The rest of the paper is organized as follows. In Section 2 we introduce and discuss necessary preliminaries such as fundamental properties of measures of risk and variability, including the notion of co-monotonicity, and we also elucidate the role of the Choquet integral in our considerations. In Section 3, starting with the classical variance and the Gini mean difference, we lay out the motivation and the origins of our Gini-based idea, and in turn introduce what we call the tail-Gini functional. In Section 4 we introduce the notion of the Gini shortfall and explore its various properties and advantages. In Section 5 we derive closed-form expressions for the Gini shortfall in the case of several parametric families of distributions, including the normal, Student-t, and more generally, elliptical distributions, as well as certain skew distributions. In Section 6 we extend our considerations to introduce a capital allocation rule, that we call the Gini shortfall allocation, and then illustrate it on a portfolio of elliptical risks. We further elucidate our general considerations in Section 7 where we analyse a portfolio of risks of a bancassurance company.

2 Preliminaries

We work with an atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let L^r denote the set of all rv's on $(\Omega, \mathcal{A}, \mathbb{P})$ with finite r-th moment, $r \in [0, \infty)$, and let L^∞ be the set of all essentially bounded rv's. Throughout the paper, positive (negative) values of $X \in L^0$ represent financial losses (profits). For every $X \in L^0$, we use F_X to denote the cdf of X, and U_X to denote any uniform $[0, 1]$ rv such that the equation $F_X^{-1}(U_X) = X$ holds almost surely. The existence of such rv's is given, for example, in Proposition 1.3 of Rüschendorf (2013). We assume that the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is rich enough so that for any set of rv's $X_1, \ldots, X_n \in L^0$ there is always a non-constant rv V independent of $X_1, \ldots, X_n$. We deal with several convex cones $\mathcal{X}$ of rv's, of which $\mathcal{X} = L^1$ is of particular importance and L^∞ is always contained in $\mathcal{X}$. We use $\mathbb{I}$ for the indicator function.

2.1 Measures of risk

For any convex cone $\mathcal{X}$ of rv's, a *risk measure* ρ is a functional that maps $\mathcal{X}$ to $(-\infty, \infty]$. Below we outline several properties that are important in the literature of risk measures, and we start with law-invariance, which is satisfied by all the risk measures that we consider.

(A) *Law-invariance*: if $X \in \mathcal{X}$ and $Y \in \mathcal{X}$ have the same distributions under $\mathbb{P}$, succinctly $X \overset{\mathrm{d}}{=} Y$, then $\rho(X) = \rho(Y)$.

The following properties have been standard in the theory of coherent risk measures (Artzner et al. (1999); also Föllmer and Schied (2002)):

(B1) *Monotonicity*: $\rho(X) \le \rho(Y)$ when $X, Y \in \mathcal{X}$ are such that $X \le Y$ $\mathbb{P}$-almost surely.

(B2) *Translation invariance*: $\rho(X - m) = \rho(X) - m$ for all $m \in \mathbb{R}$ and $X \in \mathcal{X}$.

(A1) *Positive homogeneity*: $\rho(\lambda X) = \lambda \rho(X)$ for all $\lambda > 0$ and $X \in \mathcal{X}$.

(A2) *Sub-additivity*: $\rho(X + Y) \le \rho(X) + \rho(Y)$ for all $X, Y \in \mathcal{X}$.

(A3) *Convexity*: $\rho(\lambda X + (1-\lambda)Y) \le \lambda\rho(X) + (1-\lambda)\rho(Y)$ for all $\lambda \in [0, 1]$ and $X, Y \in \mathcal{X}$.

We refer to Föllmer and Schied (2011, Chapter 4), Delbaen (2012), and McNeil et al. (2015) for interpretations of these properties. It is well known that any pair among three properties (A1), (A2) and (A3) implies the remaining one.

Definition 2.1 (Artzner et al. (1999)). A risk measure is *monetary* if it satisfies properties (B1) and (B2), and a risk measure is *coherent* if it satisfies (B1), (B2), (A1) and (A2).

Another important property of risk measures is *co-monotonic additivity*, which is based on the following notion (Schmeidler (1986)).

Definition 2.2. Two rv's X and Y are *co-monotonic* when

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0 \quad \text{for} \quad (\omega, \omega') \in \Omega \times \Omega \quad (\mathbb{P} \times \mathbb{P})\text{-almost surely.}$$

Co-monotonicity of X and Y is equivalent to the existence of a rv $Z \in L^0$ and two non-decreasing functions f and g such that $X = f(Z)$ and $Y = g(Z)$ almost surely. We refer to Dhaene et al. (2002) for an overview on co-monotonicity.

(A4) *Co-monotonic additivity*: $\rho(X + Y) = \rho(X) + \rho(Y)$ for every co-monotonic pair $X, Y \in \mathcal{X}$.

2.2 VaR, ES and the Choquet integral

We recall that the value-at-risk functional $\mathrm{VaR}_p : L^0 \to \mathbb{R}$ is defined by equation (1.1), and the corresponding expected shortfall functional $\mathrm{ES}_p : L^1 \to \mathbb{R}$ is given by equation (1.2). Obviously, when $p = 0$, then $\mathrm{ES}_0(X)$ is the average $\mathbb{E}[X]$ of X. Furthermore, both functionals VaR_p and ES_p are monetary and co-monotonically additive, whereas ES_p is also coherent. As noted in Section 1, ES_p is equal to TCE_p defined by equation (1.3) whenever the cdf F_X is continuous. For more details on various properties of these regulatory risk measures, we refer to, e.g., McNeil et al. (2015).

We next recall the Choquet integral (e.g., Denneberg (1994)) that plays a pivotal role in our following considerations. To begin with, $h : [0, 1] \to \mathbb{R}$ is called a *distortion function* when it is non-decreasing and satisfies the boundary conditions $h(0) = 0$ and $h(1) = 1$. Whenever $h : [0, 1] \to \mathbb{R}$ is of finite variation and such that $h(0) = 0$, the functional defined by the equation

$$I(X) = \int_0^\infty (h(1) - h(F_X(x)))\mathrm{d}x - \int_{-\infty}^0 h(F_X(x))\mathrm{d}x \tag{2.1}$$

for all $X \in \mathcal{X}$ is called the *signed Choquet integral*, and it is called the *Choquet integral* when h is a distortion function. When h is right-continuous, then equation (2.1) can be rewritten as

$$I(X) = \int_0^1 F_X^{-1}(t)\mathrm{d}h(t). \tag{2.2}$$

Furthermore, when h is absolutely continuous, with a function ϕ such that $\mathrm{d}h(t) = \phi(t)\mathrm{d}t$, then equation (2.2) becomes

$$I(X) = \int_0^1 F_X^{-1}(t)\phi(t)\mathrm{d}t. \tag{2.3}$$

In this case, ϕ is called the *weighting function* of the signed Choquet integral I. Equations (2.2) and (2.3) always define a signed Choquet integral, and we frequently use them in our considerations below.

The signed Choquet integral is clearly co-monotonically additive, which is readily seen from representation (2.2) (Schmeidler (1986)). Moreover, we know from Yaari (1987) and Theorem 4.88 of Föllmer and Schied (2011) that any law-invariant risk measure is co-monotonically additive and monetary if and only if it can be represented as a Choquet integral. Finally, the functional I defined by equation (2.1) is sub-additive if and only if the function h is convex (e.g. Yaari (1987) and Acerbi (2002)).

2.3 Measures of variability

Measures of variability, used to quantify the magnitude of variability of rv's, are functionals that map $\mathcal{X}$ to $[0, \infty]$. Desirable properties of measures of variability can be quite different from those of risk measures. For example, for a measure of variability ν, we as a rule require:

(C1) *Standardization*: $\nu(m) = 0$ for all $m \in \mathbb{R}$.

(C2) *Location invariance*: $\nu(X - m) = \nu(X)$ for all $m \in \mathbb{R}$ and $X \in \mathcal{X}$.

For particular applications (e.g. portfolio optimization, capital allocation, risk aggregation), we may also wish ν to satisfy convexity, sub-additivity, positive homogeneity or co-monotonic additivity, which are defined in (A1)–(A4). The following set of properties that we propose for measures of variability bear similarity with the axioms of deviation measures proposed by Rockafellar et al. (2006) and further explored in, e.g., Rockafellar et al. (2008) and Grechuk et al. (2009).

Definition 2.3. A functional $\nu : \mathcal{X} \to [0, \infty]$ is a *measure of variability* if it satisfies properties (A), (C1) and (C2). A measure of variability is *coherent* if it further satisfies (A1) and (A2).

For instance, the classical measures of variability are the variance

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2], \quad X \in L^2, \tag{2.4}$$

and the standard deviation $\mathrm{SD}_0 = \sqrt{\mathrm{Var}(X)}$. The standard deviation functional is a coherent measure of variability as it satisfies properties (C1), (C2), (A1), (A2) and (A). The variance functional satisfies properties (C1), (C2), (A) but not (A1) or (A2), and hence it is not coherent in our terminology. Note that neither the variance nor the standard deviation satisfies co-monotonic additivity (A4).

The concept of measures of variability that we propose is admittedly very similar to the deviation measures of Rockafellar et al. (2006). At the outset, we point out two differences. First, our measure of variability is law-invariant, which is a desirable property because we are interested in the distributional variation of risks. Second, and more importantly, the measures considered in this paper are not necessarily strictly-positive for all non-constant rv's, thus allowing us to focus on the variability of risks in the tail (e.g., large losses in the insurance context) while ignoring the variability (or lack of it) in the surplus. These are typical and crucial considerations when using tail-based risk measures in capital adequacy. The definition of a coherent measure of variability is the same as that of a deviation measure of Rockafellar et al. (2006) except for the above two points.

A terminological reason to introduce measures of variability is that the measures of interest in this paper (Section 3) are center-free, hence calling them "deviation" measures may not be the most accurate. Most of the mathematical results on deviation measures in Rockafellar et al. (2006, 2008) and Grechuk et al. (2009) hold for measures of variability. Below we give the characterization for co-monotonically additive coherent measures of variability, essentially established in Grechuk et al. (2009, Proposition 2.4), though in a somewhat different form.

Theorem 2.1. *For $r \in [1, \infty)$, let $\nu : L^r \to \mathbb{R}$ be any L^r-continuous functional. The following three statements are equivalent:*

(i) *ν is a co-monotonically additive and coherent measure of variability.*

(ii) *There is a convex function $h : [0, 1] \to \mathbb{R}$, $h(0) = h(1) = 0$, such that*

$$\nu(X) = \int_0^1 F_X^{-1}(u)\mathrm{d}h(u), \quad X \in L^r. \tag{2.5}$$

(iii) *There is a non-decreasing function $g : [0, 1] \to \mathbb{R}$ such that*

$$\nu(X) = \mathrm{Cov}[X, g(U_X)], \quad X \in L^r, \tag{2.6}$$

where Cov *is the covariance functional.*

Proof. For (iii)⇒(i), it is straightforward to check that equation (2.6) defines a co-monotonically additive coherent measure of variability. To show that (ii)⇒(iii), since h is almost everywhere differentiable in $[0,1]$, we can take g such that $g(t)\mathrm{d}t = \mathrm{d}h(t)$. Then

$$\nu(X) = \int_0^1 F_X^{-1}(u)\mathrm{d}h(u) = \int_0^1 F_X^{-1}(u)g(u)\mathrm{d}u = \mathbb{E}[Xg(U_X)].$$

Also note that $\mathbb{E}[g(U_X)] = \int_0^1 g(u)\mathrm{d}u = h(1) - h(0) = 0$, and so $\nu(X) = \mathrm{Cov}[X, g(U_X)]$. It remains to show (i)⇒(ii).

By Proposition 2 of Schmeidler (1986), the functional ν admits the signed Choquet integral representation

$$\nu(X) = \int_0^\infty (\xi(\Omega) - \xi(X \le x))\mathrm{d}x - \int_{-\infty}^0 \xi(X \le x)\mathrm{d}x, \quad X \in L^r,$$

where $\xi : \mathcal{F} \to \mathbb{R}$ is given by $\xi(A) = \nu(\mathbb{I}(A))$. From $\nu(1) = 0$, we have $\xi(\Omega) = 0$ and therefore

$$\nu(X) = -\int_{\mathbb{R}} \xi(X \le x)\mathrm{d}x, \quad X \in L^r.$$

Since ν is law-invariant, $\xi(X \le x)$ is a function of $\mathbb{P}(X \le x)$ and hence we can write $\xi(X \le x) = h(F_X(x))$. Note that $h(0) = 0$ and $h(1) = 0$, and

$$\nu(X) = -\int_{\mathbb{R}} h(F_X(x))\mathrm{d}x, \quad X \in L^r.$$

By Theorem 2 of De Waegenaere and Wakker (2001), convexity of ν implies that h is convex and hence right-continuous. Via equation (2.2) we arrive at

$$\nu(X) = \int_0^1 F_X^{-1}(u)\mathrm{d}h(u), \quad X \in L^r.$$

This completes the proof of Theorem 2.1. □

Theorem 2.1 provides a guideline for the appearance of co-monotonically additive coherent measures of variability: they have either representation (2.5) or (2.6). A natural choice of g in (2.6) may be the identity on $[0,1]$; this shall be discussed in Section 3 below. If one drops co-monotonic additivity, then every L^r-continuous coherent measure of variability ν has the sup-covariance representation

$$\nu(X) = \sup_{g \in \mathcal{G}} \mathrm{Cov}[X, g(U_X)], \quad X \in L^r, \tag{2.7}$$

where $\mathcal{G}$ is the set of all non-decreasing functions on $[0,1]$. For various characterization results of deviation measures, and also of coherent measures of variability, we refer to Rockafellar et al. (2006), Grechuk et al. (2009), and the references therein.

Finally, we discuss a few partial orders of variability that have been popular in economics, insurance, finance, and probability theory.

Definition 2.4. For $X, Y \in L^1$, we say that X is *second-order stochastically dominated* (SSD) by Y, succinctly $X \preceq_{\mathrm{SSD}} Y$, if $\mathbb{E}[f(X)] \le \mathbb{E}[f(Y)]$ for all increasing convex functions f, assuming that both expectations exist. If, in addition, $\mathbb{E}[X] = \mathbb{E}[Y]$, then we say that X is *smaller than Y in convex order*, succinctly $X \preceq_{\mathrm{CX}} Y$.

Both SSD and CX orders describe dominance in terms of variability. Since a measure of variability ν is always standardized, it is often desirable for a measure of variability ν to be monotone with respect to CX. Similarly, if a risk measure ρ is obtained by combining a measure of variability and another risk measure, then it may be desirable for ρ to be monotone with respect to SSD. Hence, it is natural to introduce the following two properties:

(B3) *SSD-monotonicity*: if $X \preceq_{\mathrm{SSD}} Y$, then $\rho(X) \le \rho(Y)$.

(C3) *CX-monotonicity*: if $X \preceq_{\mathrm{CX}} Y$, then $\nu(X) \le \nu(Y)$.

It is clear that for the same functional, i.e. $\rho = \nu$, property (C3) is weaker than (B3) since $\preceq_{\mathrm{CX}}$ implies $\preceq_{\mathrm{SSD}}$. The standard deviation and variance functionals on L^2 satisfy property (C3). The expected shortfall ES_p satisfies (B3) for every $p \in (0, 1)$. In fact, on L^q, $q \in [1, \infty]$, all real-valued coherent measures of variability are CX-monotone, and all real-valued law-invariant coherent risk measures are SSD-monotone. We refer to Dana (2005), Grechuk et al. (2009), and Föllmer and Schied (2011) for proofs of the above assertions, and to Mao and Wang (2016) for a characterization of SSD-monotone risk measures.

Remark 2.1. In this section, several properties of measures of risk and variability are presented. Speaking generally, whether specific properties such as convexity, sub-additivity, positive homogeneity or co-monotonic additivity are reasonable/desirable or not depends on the underlying application[2]. For instance, in the context of portfolio selection, convexity is a natural property to consider, whereas in the context of risk aggregation or risk capital allocation, sub-additivity is common. Desirable properties in the context of financial regulation are discussed in several recent papers (e.g., Embrechts et al. (2014), Emmer et al. (2015) and Föllmer and Weber (2015)).

3 Classical and tail-based Gini functionals

Throughout the rest of this paper, unless explicitly noted otherwise, we work with the cone $\mathcal{X} = L^1$ as the natural domain of our measures of risk and variability.

[2]We are grateful to an anonymous referee for raising this point.

3.1 Classical Gini functional and the signed Choquet integral

Our main idea of this paper originates from the work of Corrado Gini who argued more than a hundred years ago (e.g., Giorgi (1990, 1993) and Ceriani and Verme (2012) for references and historical notes) that representation (2.4) of the variance $\mathrm{Var}(X)$ might be misleading in the sense that variability of any rv should not be based on the center of the corresponding distribution. Consequently, C. Gini noted the following alternative expression for the variance

$$\mathrm{Var}(X) = \frac{1}{2}\mathbb{E}[(X^* - X^{**})^2], \quad X \in L^2,$$

where X^* and X^{**} are two independent copies of X. This representation is free of any center, but it raises a further question about the rationale of using the quadratic function $(x-y)^2$ because it distorts the values of $X^* - X^{**}$ by making them larger when they are outside the interval $[-1, 1]$ and smaller otherwise. Even the square root in the definition of standard deviation does not rectify the problem, as we have already argued in the context of the TSD risk measure in Section 1. This reasoning led C. Gini to the idea of introducing the variability measure

$$\mathrm{Gini}(X) = \mathbb{E}[|X^* - X^{**}|], \quad X \in L^1, \tag{3.1}$$

which is nowadays known as the Gini mean difference; we call $\mathrm{Gini} : L^1 \to [0, \infty)$ the Gini functional throughout this paper. The Gini functional has been remarkably influential in numerous research areas, applied and theoretical (e.g., Yitzhaki and Schechtman (2013), and the references therein). Note that definition (3.1) can be rewritten as

$$\mathrm{Gini}(X) = \int_0^1 \int_0^1 |F_X^{-1}(u) - F_X^{-1}(v)| \mathrm{d}u \mathrm{d}v. \tag{3.2}$$

Our next step in developing the main idea of the present paper is based on the observation of Denneberg (1990) that the Gini functional is co-monotonically additive, that is, the equation $\mathrm{Gini}(X + Y) = \mathrm{Gini}(X) + \mathrm{Gini}(Y)$ holds for every co-monotonic pair X and Y in L^1. The co-monotonic additivity of the Gini functional follows immediately from the fact that it is a signed Choquet integral, that is, the representation

$$\mathrm{Gini}(X) = 2 \int_0^1 F_X^{-1}(u)(2u - 1)\mathrm{d}u \tag{3.3}$$

holds for every $X \in L^1$. Equation (3.3) is of course well-known (it is also a special case of Proposition 3.2 below). The next corollary follows immediately from Theorem 2.1 and the fact that all continuous coherent measures of variability are CX-monotone.

Corollary 3.1. *The Gini functional is a coherent measure of variability, and it is* CX*-monotone.*

It is a simple exercise to check that equation (3.3) can be rewritten as a covariance:

$$\mathrm{Gini}(X) = 4\mathrm{Cov}[F_X^{-1}(U), U], \tag{3.4}$$

where U can be any uniformly on $[0, 1]$ distributed rv. This interpretation of the Gini functional provides a pivotal starting point for constructing Gini-based risk measures and capital allocation rules that we introduce and explore later. Equation (3.4) can further be written as

$$\mathrm{Gini}(X) = 4\mathrm{Cov}[X, U_X]; \tag{3.5}$$

recall that U_X is a uniform $[0, 1]$ rv such that the equation $F_X^{-1}(U_X) = X$ holds almost surely. In the spirit of Theorem 2.1(iii), the Gini functional offers a most natural choice of co-monotonically additive and coherent measure of variability, where g in (2.6) is chosen as the identity on $[0, 1]$.

We conclude this subsection with the note that closed-form expressions for the Gini functional and related quantities in the case of many economic-size distributions can be found in numerous articles and books dealing with measures of economic inequality. In addition, in Section 5 below we provide closed-form expressions for the Gini functional for several parametric families of interest in financial and actuarial risk modelling.

3.2 Tail-Gini functional

In the modern 'prudent' financial risk management, practitioners and researchers often look at the tail risk. The value-at-risk and the expected shortfall (Section 2.2 above) are risk measures that conform to such philosophy, but none of them appropriately reflects tail variability. Therefore, we next introduce the tail-Gini functional (TGini).

Given any risk rv $X \in L^1$ and a prudence level $p \in [0, 1)$, let $F_{X,p}$ denote the cdf of the rv $F_X^{-1}(U_p)$, where U_p is uniformly distributed on $[p, 1]$. Then the tail-Gini functional is given by

$$\mathrm{TGini}_p(X) = \mathbb{E}[|X_p^* - X_p^{**}|\,], \tag{3.6}$$

where the rv's X_p^* and X_p^{**} are two independent copies with the cdf $F_{X,p}$. Obviously, when $p = 0$, then $\mathrm{TGini}_0(X)$ is equal to $\mathrm{Gini}(X)$, as is easily seen either directly from equation (3.6) or by comparing equation (3.2) and the following representation of the TGini functional

$$\mathrm{TGini}_p(X) = \frac{1}{(1-p)^2} \int_p^1 \int_p^1 |F_X^{-1}(u) - F_X^{-1}(v)| \mathrm{d}u \mathrm{d}v, \quad X \in L^1. \tag{3.7}$$

To work out additional intuition, assume for a moment that the cdf F_X is continuous. Then the tail-Gini functional can be written in the form of a conditional covariance:

$$\mathrm{TGini}_p(X) = \frac{4}{1-p}\mathrm{Cov}[X, F_X(X) \mid X > x_p]. \tag{3.8}$$

Alternatively, the functional can be written in the form of a conditional expectation:

$$\mathrm{TGini}_p(X) = \mathbb{E}[|X^* - X^{**}| \mid X^* > x_p, X^{**} > x_p], \tag{3.9}$$

where X^* and X^{**} are two independent copies of X. Setting $p = 0$ reduces equation (3.8) to formula (3.4) for Gini(X), and equation (3.9) to original Gini definition (3.1).

Just like the Gini functional, for any (continuous or not) cdf F_X, the tail-Gini functional can be represented as a signed Choquet integral.

Proposition 3.2. *For every $p \in (0,1)$, the tail-Gini functional is a signed Choquet integral given by the equation*

$$\mathrm{TGini}_p(X) = \frac{2}{(1-p)^2}\int_p^1 F_X^{-1}(u)(2u - (1+p))\mathrm{d}u. \tag{3.10}$$

Therefore, the tail-Gini functional is co-monotonically additive.

Proof. Direct calculations give

$$\begin{aligned}
\mathrm{TGini}_p(X) &= \frac{2}{(1-p)^2}\int_p^1 \left(\int_u^1 (F_X^{-1}(v) - F_X^{-1}(u))\mathrm{d}v\right)\mathrm{d}u \\
&= \frac{2}{(1-p)^2}\int_p^1 \left(\int_u^1 F_X^{-1}(v)\mathrm{d}v - F_X^{-1}(u)(1-u)\right)\mathrm{d}u \\
&= \frac{2}{(1-p)^2}\left(\int_p^1 F_X^{-1}(v)(v-p)\mathrm{d}v - \int_p^1 F_X^{-1}(u)(1-u)\mathrm{d}u\right) \\
&= \frac{2}{(1-p)^2}\int_p^1 F_X^{-1}(u)\big(2u - (1+p)\big)\mathrm{d}u.
\end{aligned}$$

This completes the proof of Proposition 3.2. □

Similarly to the Gini functional, it is easy to see that the tail-Gini functional is law-invariant, standardized, location invariant, and positively homogeneous. However, the tail-Gini functional is not sub-additive for any $p \in (0,1)$, as shown in Proposition 3.3 below (one may also directly compare (2.5) and (3.10)). Therefore, unlike the Gini functional, the tail-Gini functional is not a coherent measure of variability.

Proposition 3.3. *For every $p \in (0,1)$, the tail-Gini functional* TGini_p *is not sub-additive.*

Proof. Note first that $\mathrm{TGini}_p(c) = 0$ for every constant $c \in \mathbb{R}$, and $\mathrm{TGini}_p(X) > 0$ for every rv X such that F_X^{-1} is not constant over the interval $(p, 1)$. Let X be such that $\mathbb{P}(X = -1) = p$ and $\mathbb{P}(X = 0) = 1 - p$, and let the rv's X and Y be independent and identically distributed. Obviously $\mathbb{P}(X+Y = 0) = (1-p)^2 < 1-p$, which means that the quantile function F_{X+Y}^{-1} is not constant over the interval $(p, 1)$. From the above arguments we have

$$\mathrm{TGini}_p(X + Y) > 0 = 0 + 0 = \mathrm{TGini}_p(X) + \mathrm{TGini}_p(Y).$$

Moreover, for every $p \in (0, 1)$, the functional TGini_p is not CX-monotone, which can be seen from the fact that $X + Y \preceq_{\mathrm{CX}} 2X$ and

$$\mathrm{TGini}_p(X + Y) > 0 = \mathrm{TGini}_p(2X).$$

This completes the proof of Proposition 3.3. □

From the proof of Proposition 3.3, we see that $\mathrm{TGini}_p(X)$ may be zero even if the rv X is not constant. This property violates the definition of deviation measures in Rockafellar et al. (2006), but it is essential to any tail-based measure of variability. Although TGini_p is not a coherent measure of variability, we see in the next section that when combined with ES_p, it gives rise to a coherent risk measure that quantifies both the magnitude and the variability of tail risks.

4 Gini shortfall

Here we introduce the Gini shortfall (GS), which is a linear combination of the expected shortfall ES_p and the tail-Gini functional TGini_p. Namely,

$$\mathrm{GS}_p^\lambda(X) = \mathrm{ES}_p(X) + \lambda\,\mathrm{TGini}_p(X), \quad X \in L^1, \tag{4.1}$$

where $p \in [0, 1)$ is the prudence level and $\lambda \geq 0$ is the loading parameter. GS yields a two-parameter class of tail risk measures in the sense of Liu and Wang (2016). The case $p = 0$ needs to be considered separately because of mathematical and terminological reasons. A mathematical reason will be given in Remark 4.2 below, once the necessarily background has been established. For a terminological reason, we note that, for any $\lambda \geq 0$, the functional

$$\mathrm{GS}_0^\lambda(X) = \mathbb{E}[X] + \lambda\,\mathrm{Gini}(X), \quad X \in L^1, \tag{4.2}$$

was originally introduced (under a different notation) by Denneberg (1990) and called the Gini principle (see Remark 4.2).

4.1 Basic properties

We let $p \in (0,1)$ throughout this subsection, unless explicitly noted otherwise. For the functional GS_p^λ to be a reasonable risk measure, it should satisfy some desirable properties listed in Section 2.1. Specifically, in the previous section we noted that, for $p \in (0,1)$, the functional TGini_p is not sub-additive and, as a measure of variability, it is not monotone. Therefore, in order to make GS_p^λ monotone or sub-additive, the parameter $\lambda \geq 0$ cannot be too large. Indeed, when λ is zero, then GS_p^λ obviously inherits all the properties of the expected shortfall ES_p, but when λ is sufficiently large, then the TGini_p-term starts to dominate ES_p, and thus monotonicity and sub-additivity of GS_p^λ cannot be expected to hold. This suggests that there might be a threshold that delineates the values of $\lambda \in (0,\infty)$ for which GS_p^λ is monotone and/or sub-additive. As we show in the next theorem, the thresholds for both monotonicity and sub-additivity are the same, and equal to $1/2$.

Theorem 4.1. *Let $p \in (0,1)$ and $\lambda \in [0,\infty)$.*

(1) *The Gini shortfall GS_p^λ is a signed Choquet integral given by the equation*

$$\mathrm{GS}_p^\lambda(X) = \int_0^1 F_X^{-1}(u)\phi_{p,\lambda}(u)\mathrm{d}u, \tag{4.3}$$

with the function

$$\phi_{p,\lambda}(u) = \frac{1}{(1-p)^2}\left(1-p+4\lambda\left(u-\frac{1+p}{2}\right)\right)\mathbb{I}_{[p,1]}(u), \quad u \in [0,1], \tag{4.4}$$

where $\mathbb{I}_{[p,1]}$ is the indicator function of the interval $[p,1]$.

(2) *The functional GS_p^λ is translation invariant, positively homogeneous, and co-monotonically additive.*

(3) *The following statements are equivalent:*

(i) GS_p^λ *is monotone;*

(ii) GS_p^λ *is sub-additive;*

(iii) GS_p^λ *is SSD-monotone;*

(iv) GS_p^λ *is a coherent risk measure;*

(v) $\lambda \in [0,1/2]$.

Proof. Part (1) follows from the equations:

$$\begin{aligned}
\mathrm{GS}_p^\lambda(X) &= \mathrm{ES}_p(X) + \lambda\,\mathrm{TGini}_p(X) \\
&= \frac{1}{1-p}\int_p^1 F_X^{-1}(u)\mathrm{d}u + \frac{4\lambda}{(1-p)^2}\int_p^1 F_X^{-1}(u)\left(u - \frac{1+p}{2}\right)\mathrm{d}u \\
&= \frac{1}{(1-p)^2}\int_p^1 F_X^{-1}(u)\left(1-p+4\lambda\left(u-\frac{1+p}{2}\right)\right)\mathrm{d}u \\
&= \int_0^1 F_X^{-1}(u)\phi_{p,\lambda}(u)\mathrm{d}u.
\end{aligned}$$

To prove part (2), we note that co-monotonic additivity and positive homogeneity arise directly from equation (4.3). For translation invariance, we note that $\mathrm{TGini}_p(X+c) = \mathrm{TGini}_p(X)$ for all $c \in \mathbb{R}$, and so

$$\begin{aligned}
\mathrm{GS}_p^\lambda(X+c) &= \mathrm{ES}_p(X+c) + \lambda\mathrm{TGini}_p(X+c) \\
&= c + \mathrm{ES}_p(X) + \lambda\mathrm{TGini}_p(X).
\end{aligned}$$

To prove part (3), we need an auxiliary result, which we formulate as Lemma 4.2 below. Noting that $\phi_{p,\lambda}(u) = 0$ for all $u \in [0,p)$, and that $\phi_{p,\lambda}$ is an increasing function on $[p,1]$, elementary analysis shows that $\phi_{p,\lambda}$ is non-negative if and only if $\lambda \in [0,1/2]$ and, moreover, $\phi_{p,\lambda}$ is non-decreasing if and only if $\lambda \in [0,1/2]$. Lemma 4.2 implies that statements (i), (ii) and (v) are equivalent. The equivalence (iv)$\Leftrightarrow$(i)+(ii) is trivial because GS_p^λ is translation invariant and positively homogeneous. Note that statement (iv) implies (iii) (e.g., Corollary 4.65 in Föllmer and Schied (2011)), which in turn implies statement (i). This proves that all statements (i)–(v) are equivalent, and thus completes the proof of Theorem 4.1. □

In the following lemmas, we say that a function $\phi \in L^\infty([0,1])$ is a.e. non-decreasing if for all $a,b \in [0,1]$, $a<b$ and $\epsilon \in (0, a-b)$, it holds that $\int_a^{a+\epsilon}\phi(u)\mathrm{d}u \le \int_{b-\epsilon}^b \phi(u)\mathrm{d}u$.

Lemma 4.2. *For $\phi \in L^\infty([0,1])$, let the functional $R_\phi : L^1 \to \mathbb{R}$ be defined by the equation*

$$R_\phi(X) = \int_0^1 F_X^{-1}(u)\phi(u)\mathrm{d}u. \tag{4.5}$$

The following statements hold:

(a) *R_ϕ is monotone if and only if $\phi \ge 0$ on $[0,1]$ a.e.;*

(b) *R_ϕ is sub-additive if and only if ϕ is non-decreasing on $[0,1]$ a.e.*

Below we quote Theorem 4.1 of Acerbi (2002), which is similar to and slightly weaker than Lemma 4.2.

Lemma 4.3 (Theorem 4.1 of Acerbi (2002), adjusted to our sign convention)**.** *For* $\phi \in L^\infty([0,1])$*, let the functional* $R_\phi : L^1 \to \mathbb{R}$ *be defined by equation* (4.5)*. The following are equivalent:*

(i) R_ϕ *is monotone, sub-additive and translation invariant.*

(ii) $\int_0^1 \phi(u)\mathrm{d}u = 1$*, and* $\phi \geq 0$ *and* ϕ *is non-decreasing on* $[0,1]$ *a.e.*

Furthermore, two counter-examples in the proof of Theorem 4.1 of Acerbi (2002) reveal that R_ϕ is monotone only if $\phi \geq 0$ on $[0,1]$ a.e., and R_ϕ is sub-additive only if ϕ is non-decreasing on $[0,1]$ a.e. Hence, to show Lemma 4.2, it remains to show the "if" direction of both (a) and (b).

Proof of Lemma 4.2. For part (a), note that when $X \leq Y$, then $F_X^{-1}(u) \leq F_Y^{-1}(u)$ for all $u \in [0,1]$. Using this fact together with the non-negativity of ϕ, we obtain $R_\phi(Y) \leq R_\phi(X)$. For part (b), suppose that ϕ is non-decreasing. Let $M = \text{ess-inf}_{u\in[0,1]}\phi(u)$, which is finite because $\phi \in L^\infty([0,1])$. Let $\phi_+(u) = M + \phi(u)$, $u \in [0,1]$. Noting that ϕ_+ is non-negative and non-decreasing, by Theorem 4.1 of Acerbi (2002) we have that the functional $R_{\phi_+}/\|\phi_+\|_1$ is a coherent risk measure and hence sub-additive. As the functional $R_{\phi_+} - R_\phi$, which is equal to $M\mathbb{E}[\cdot]$, is additive, we have that R_ϕ is sub-additive. □

Remark 4.1. One can show that the functional R_ϕ is consistent with SSD if and only if both statements (a) and (b) hold. This result has essentially been obtained by Yaari (1987) although the formulation in the noted paper is different from ours. From this result, the equivalence of statements (i)–(v) in Theorem 4.1 becomes clear without necessarily consulting Föllmer and Schied (2011).

Remark 4.2. The Gini shortfall GS_0^λ, which is called the Gini principle by Denneberg (1990), has not been included into Theorem 4.1 due to the following mathematical reason. Namely, from Corollary 3.1 and Lemma 4.2, the functional GS_0^λ is always translation invariant, positively homogeneous, co-monotonically additive, and sub-additive. This is in stark contrast with the case $p \in (0,1)$: the functional GS_p^λ is sub-additive only when $\lambda \in [0,1/2]$.

4.2 Continuity properties

In this section we study continuity properties of the Gini shortfall GS_p^λ with respect to certain types of convergence. Since no distinction between the cases $p = 0$ and $p \in (0,1)$ is necessary, throughout the section we work with $p \in [0,1)$.

The continuity of law-invariant risk measures corresponds to Hampel's classic notion of *qualitative robustness*, which has been a focal point in the recent study of risk measures (e.g., Krätschmer et al. (2014), Embrechts et al. (2014, 2015), and Föllmer and Weber (2015)). Given that we work with Gini-type functionals, it is not surprising that the Wasserstein distance becomes a natural tool in our context: for two rv's X and Y, the distance is defined by (Dobrushin (1970))

$$W_1(X,Y) = \sup \mathbb{E}[|X^* - Y^*|]$$

with the supremum taken over all rv's $X^* \sim F_X$ and $Y^* \sim F_Y$. The Wasserstein distance can equivalently be written as (Dobrushin (1970))

$$W_1(X,Y) = \int_0^1 |F_X^{-1}(u) - F_Y^{-1}(u)|\mathrm{d}u.$$

Theorem 4.4. *For $p \in [0,1)$ and $\lambda \in [0,\infty)$, the following statements hold:*

(i) GS_p^λ *is continuous with respect to the Wasserstein distance in L^1;*

(ii) GS_p^λ *is continuous with respect to the L^1-norm;*

(iii) *for every $M > 0$, the functional GS_p^λ is continuous with respect to weak convergence in the subspace $L_M = \{X \in L^1 : |X| \le M\}$ of L^1.*

Proof. For the function $\phi_{p,\lambda}$ defined by equation (4.4), we have

$$\begin{aligned}|\phi_{p,\lambda}(u)| &\le \max\left\{\frac{1-p+2\lambda(1-p)}{(1-p)^2}, \frac{|1-p-2\lambda(1-p)|}{(1-p)^2}\right\}\\ &\le \frac{1+2\lambda}{1-p} =: c_{p,\lambda} < \infty.\end{aligned}$$

Hence,

$$\begin{aligned}|\mathrm{GS}_p^\lambda(X_n) - \mathrm{GS}_p^\lambda(X)| &\le \int_0^1 |F_{X_n}^{-1}(u)\phi(u) - F_X^{-1}(u)\phi_{p,\lambda}(u)|\mathrm{d}u\\ &\le c_{p,\lambda}\int_0^1 |F_{X_n}^{-1}(u) - F_X^{-1}(u)|\mathrm{d}u. \qquad (4.6)\end{aligned}$$

Part (i) follows from bound (4.6) because if $X_n \to X$ in the Wasserstein distance, then

$$|\mathrm{GS}_p^\lambda(X_n) - \mathrm{GS}_p^\lambda(X)| \le c_{p,\lambda} W_1(X_n, X) \to 0$$

when $n \to \infty$. Part (ii) follows from part (i) because $\mathbb{E}[|X_n - X|] \le W_1(X_n, X)$. To prove part (iii), note that when $X_1, X_2, \ldots \in L_M$ and $X_n \to X \in L_M$ weakly, then

$F_{X_n}^{-1}(t) \to F_X^{-1}(t)$ for all continuity points $t \in [0,1]$ of the quantile function F_X^{-1}. Bound (4.6) together with the Bounded Convergence Theorem imply $|\mathrm{GS}_p^\lambda(X_n) - \mathrm{GS}_p^\lambda(X)| \to 0$ when $n \to \infty$. This establishes the continuity of GS_p^λ with respect to weak convergence in L_M, thus completing the proof of Theorem 4.4. □

We conclude this section with a few additional observations regarding the continuity of the Gini shortfall.

Remark 4.3. From the proof of Theorem 4.4 we see that for a signed Choquet integral to have continuity properties (i)–(iii), it is sufficient to have a bounded weighting function. Because of this reason, the functionals ES_p and TGini_p satisfy the three continuity properties. For more results on the continuity properties of distortion risk measures, we refer to Emmer et al. (2015), and Föllmer and Weber (2015).

Remark 4.4. Since the Gini shortfall is continuous with respect to the L^1-metric, it is also continuous with respect to any stronger metric, such as the L^2- and L^∞-metrics.

Remark 4.5. Another way to establish statement (ii) is to use Corollary 2.3 of Kaina and Rüschendorf (2009), which says that a finite-valued convex risk measure on L^1 is continuous with respect to the L^1-norm.

Remark 4.6. By Theorem 2.4 of Embrechts et al. (2015), the functional GS_p^λ is *aggregation-robust*, which means that $\mathrm{GS}_p^\lambda(X_1 + \cdots + X_n)$ is continuous with respect to convergence in the dependence structure (copula) of $(X_1, \ldots, X_n)$, assuming that the marginal distributions are fixed.

4.3 Comparison of tail variability

In this section, to further study features of the Gini shortfall, we introduce an ordering of tail variability, similarly to the partial orders of variability in Definition 2.4 but with a focus on the tail distribution. Recall that for any rv $X \in L^1$ and $p \in [0,1)$, we denote by $F_{X,p}$ the cdf of the rv $F_X^{-1}(U_p)$, where U_p is uniformly distributed on $[p,1]$.

Definition 4.1. For $X, Y \in L^1$, we say that Y has a *larger p-tail variability compared to X*, succinctly $X \preceq_{p\text{-CX}} Y$, if $F_X^{-1}(U_p) \preceq_{\mathrm{CX}} F_Y^{-1}(U_p)$. If, in addition, $F_{X,p}$ and $F_{Y,p}$ are not identical, then we say that Y has a *strictly larger* p-tail variability compared to X, succinctly $X \prec_{p\text{-CX}} Y$.

Intuitively, the partial order $\preceq_{p\text{-CX}}$ compares the variability of the two tail distributions $F_{X,p}$ and $F_{Y,p}$, that is, the variability of risks beyond the prudence level p. The following theorem states that the tail Gini functional is strictly monotone with respect to tail variability.

Theorem 4.5. *For* $p \in [0,1)$*,* $\lambda \in [0,\infty)$ *and* $X, Y \in L^1$*, if* $X \preceq_{p\text{-CX}} Y$*, then* $\mathrm{TGini}_p(X) \leq \mathrm{TGini}_p(Y)$*. Moreover, if* $X \prec_{p\text{-CX}} Y$*, then* $\mathrm{TGini}_p(X) < \mathrm{TGini}_p(Y)$*.*

Proof. Write $X_p = F_X^{-1}(U_p)$ and $Y_p = F_Y^{-1}(U_p)$, where U_p is uniformly distributed on $[p,1]$. Since the Gini functional is CX-monotone (Corollary 3.1) and $X_p \preceq_{\mathrm{CX}} Y_p$, we have $\mathrm{Gini}(X_p) \leq \mathrm{Gini}(Y_p)$; thus $\mathrm{TGini}_p(X) \leq \mathrm{TGini}_p(Y)$.

Now suppose $X \prec_{p\text{-CX}} Y$, which implies $\mathrm{ES}_q(X_p) \leq \mathrm{ES}_q(Y_p)$ for all $q \in (0,1)$. By Proposition 3.2,

$$
\begin{aligned}
\mathrm{TGini}_p(X) &= \frac{2}{(1-p)^2}\int_p^1 F_X^{-1}(u)(2u-(1+p))\mathrm{d}u \\
&= \frac{2}{(1-p)^2}\int_p^1 F_X^{-1}(u)\left(\int_p^u 2\mathrm{d}v-(1-p)\right)\mathrm{d}u \\
&= \frac{4}{(1-p)^2}\int_p^1\int_p^u F_X^{-1}(u)\mathrm{d}v\mathrm{d}u - \frac{2}{1-p}\int_p^1 F_X^{-1}(u)\mathrm{d}u \\
&= \frac{4}{(1-p)^2}\int_p^1\int_v^1 F_X^{-1}(u)\mathrm{d}u\mathrm{d}v - 2\mathrm{ES}_p(X) \\
&= \frac{4}{(1-p)^2}\int_p^1 (1-v)\mathrm{ES}_v(X)\mathrm{d}v - 2\mathbb{E}[X_p].
\end{aligned}
$$

As $F_{X,p}$ and $F_{Y,p}$ are not identical, there exists $q \in (0,1)$ such that $\mathrm{ES}_q(X_p) < \mathrm{ES}_q(Y_p)$. Note that, by definition, for any random variable $X \in L^1$, $\mathrm{ES}_q(X)$ is continuous with respect to q. It follows that there exists a neighborhood $[q-\epsilon, q+\epsilon]$ of q such that $\mathrm{ES}_v(X_p) < \mathrm{ES}_v(Y_p)$ for $v \in [q-\epsilon, q+\epsilon]$. Therefore, $\int_p^1(1-v)\mathrm{ES}_v(X) < \int_p^1(1-v)\mathrm{ES}_v(Y)$. Noting $\mathbb{E}[X_p] = \mathbb{E}[Y_p]$, we conclude that $\mathrm{TGini}_p(X) < \mathrm{TGini}_p(Y)$. □

For $X \in L^1$, $\mathrm{ES}_p(X)$ is the mean of the distribution $F_{X,p}$, and as such, $\mathrm{ES}_p(X) = \mathrm{ES}_p(Y)$ if either $X \preceq_{p\text{-CX}} Y$ or $Y \preceq_{p\text{-CX}} X$. From there, it is clear that the expected shortfall ES_p is not strictly monotone with respect to tail variability, whereas the Gini shortfall GS_p^λ for $\lambda > 0$ is, since GS_p^λ is a combination of ES_p and TGini_p.

Remark 4.7. In the literature, differentiation between riskiness and variability is rather vague. Indeed, some classic articles (e.g. Rothschild and Stiglitz (1970)) treat the terms "riskier" and "more variable" identically. In our approach, the Gini shortfall is a combination of an expected shortfall (a risk measure) and a tail Gini functional (a variability measure), thus capturing both the magnitude of the risk and the variability of the risk.

4.4 Comparing the Gini shortfall with other risk measures

In this section, we discuss some advantages of the new class of risk measures, the Gini shortfall. The Gini shortfall GS_p^λ is a co-monotonically additive monetary risk measure, and it is coherent for $\lambda \in [0, 1/2]$ (co-monotonically additive and coherent risk measures are known as *spectral risk measures*). When compared to other spectral risk measures, GS has the following advantages:

1. GS admits a formulation in terms of the tail risks, i.e.,

$$\mathrm{GS}_p^\lambda(X) = \mathbb{E}[X_p^*] + \lambda \mathbb{E}[|X_p^* - X_p^{**}|], \tag{4.7}$$

   where the rv's X_p^* and X_p^{**} are two independent copies having the cdf $F_{X,p}$. Form (4.7) provides remarkable tractability to the Gini shortfall (e.g. Monte-Carlo simulation), and distinguishes it from a majority of spectral risk measures.

2. GS adequately complements existing tail risk measures, especially ES and VaR, by taking into account both the tail expectation and tail variability of the underlying risks.

3. Some classes of spectral risk measures require the finiteness of higher-order moments of the underlying risks (e.g. power spectral risk measures of Dowd et al. (2008)), whereas for the Gini shortfall the finiteness of the first moment suffices. On a related note, robustness properties of GS are generally attractive (Theorem 4.4).

4. GS yields a flexible two-parameter class of tail risk measures, which is suitable for the regulatory consideration of tail risk (see, e.g., Liu and Wang (2016)).

5. GS is closely related to the Gini functional. The Gini functional is one of the most natural and well studied co-monotonically additive measures of variability. It has a clear economic interpretation, and is widely applied in many disciplines.

6. GS is strictly monotone with respect to tail variability (Section 4.3).

We note in passing that among the merits mentioned above, ES also enjoys the advantages in points 1, 3 and 4, but not the ones in points 2, 5 and 6.

5 Gini shortfall for some parametric risks

To simplify our considerations, we work with standardized rv's, collectively denoted by Z, whose location and scale parameters are 0 and 1, respectively. The general case follows immediately because when $X \stackrel{d}{=} \alpha+\beta Z$ for $\alpha \in \mathbb{R}$ and $\beta \in (0,\infty)$, then for every $p \in [0,1)$ we have

$$\mathrm{ES}_p(X) = \alpha + \beta\ \mathrm{ES}_p(Z) \tag{5.1}$$

and

$$\mathrm{TGini}_p(X) = \beta\ \mathrm{TGini}_p(Z). \tag{5.2}$$

In what follows, therefore, we concentrate on deriving closed-form expressions for $\mathrm{ES}_p(Z)$ and $\mathrm{TGini}_p(Z)$. Specifically, we start with the general elliptical family and then specialize the obtained results to normal and Student-t families that have been popular when modeling financial returns (e.g., Knight and Satchel (2001)). Then we proceed to discuss how $\mathrm{ES}_p(Z)$ and $\mathrm{TGini}_p(Z)$ can be calculated for the skew-normal and skew-t distributions (e.g. Azzalini (1985), Azzalini and Capitanio (2003)).

5.1 General formulas for elliptical risks

Let Z be a spherical rv with characteristic generator $\psi : [0,\infty) \to \mathbb{R}$; succinctly $Z \sim S(\psi)$. When Z has a probability density function (pdf), which is the case that we are interested in, then there is a density generator $g : [0,\infty) \to [0,\infty)$ such that $\int_0^\infty z^{-1/2} g(z)\mathrm{d}z < \infty$, and hence we succinctly write $Z \sim S(g)$. The pdf $f : \mathbb{R} \to [0,\infty)$ of Z can be expressed by the formula

$$f(z) = c\, g(z^2/2),$$

where $c > 0$ is the normalizing constant. The mean $\mathbb{E}[Z]$ is finite when

$$\int_0^\infty g(z)\mathrm{d}z < \infty, \tag{5.3}$$

in which case we have $\mathbb{E}[Z] = 0$ because the pdf f is symmetric around 0. Under condition (5.3), the function $\overline{G} : [0,\infty) \to [0,\infty)$ given by

$$\overline{G}(y) = c \int_y^\infty g(x)\mathrm{d}x$$

is well defined and called the tail generator of Z (see e.g. Furman and Landsman (2006a)). The function $\overline{G}$ plays a crucial role in our following considerations. Denote the p-quantile of Z by z_p.

Theorem 5.1. *When $Z \sim S(g)$ and the mean $\mathbb{E}[Z]$ is finite, then for every $p \in (0,1)$ we have*

$$\mathrm{ES}_p(Z) = \frac{\overline{G}(z_p^2/2)}{1-p} \tag{5.4}$$

and

$$\mathrm{TGini}_p(Z) = \frac{4}{1-p}\mathbb{E}\left[\overline{G}(Z^2/2) \mid Z > z_p\right] - 2\mathrm{ES}_p(Z). \tag{5.5}$$

Letting $p \downarrow 0$*, equation* (5.5) *reduces to*

$$\mathrm{Gini}(Z) = 4\mathbb{E}\left[\overline{G}(Z^2/2)\right]. \tag{5.6}$$

Proof. Equation (5.4), which is well known (e.g. Landsman and Valdez (2003), Furman and Landsman (2006a)), can easily be established by using the definition of the pdf f of Z and then appropriately changing the variable of integration. To establish equation (5.5), we first note that $F_Z(Z)$ is a uniform on $[0,1]$ rv, and thus

$$\mathrm{Cov}[Z, F(Z) \mid Z > z_p] = \frac{1}{1-p}\int_{z_p}^{\infty} zF(z)f(z)\mathrm{d}z - \frac{1+p}{2}\mathrm{ES}_p(Z). \tag{5.7}$$

Next, we use the equation $zf(z)\mathrm{d}z = -d\overline{G}(z^2/2)$, integrate by parts, and arrive at

$$\begin{aligned}\int_{z_p}^{\infty} zF(z)f(z)\mathrm{d}z &= \int_{z_p}^{\infty} \overline{G}(z^2/2)f(z)\mathrm{d}z + p\overline{G}(z_p^2/2) \\ &= \int_{z_p}^{\infty} \overline{G}(z^2/2)f(z)\mathrm{d}z + p(1-p)\mathrm{ES}_p(Z).\end{aligned} \tag{5.8}$$

Using equation (5.8) on the right-hand side of equation (5.7), we obtain

$$\mathrm{Cov}[Z, F(Z) \mid Z > z_p] = \mathbb{E}\left[\overline{G}(Z^2/2 \mid Z > z_p)\right] - \frac{1-p}{2}\mathrm{ES}_p[Z]. \tag{5.9}$$

Upon recalling representation (3.8) of the tail-Gini functional, equation (5.9) implies equation (5.5) from which equation (5.6) follows immediately. This completes the proof of Theorem 5.1. □

Remark 5.1. The variance $\mathrm{Var}(Z)$ is finite whenever $\int_0^\infty z^{1/2}g(z)\mathrm{d}z < \infty$, in which case $\mathrm{Var}(Z)$ is equal to $\int_{-\infty}^{\infty}\overline{G}(z^2/2)\mathrm{d}z$. Hence, $f^*(z) = \overline{G}(z^2/2)/\mathrm{Var}(Z)$ is a pdf. With Z^* denoting a rv that has this pdf, equation (5.5) becomes

$$\mathrm{TGini}_p(Z) = \frac{4\mathrm{Var}(Z)}{1-p}\mathbb{E}[f^*(Z) \mid Z > z_p] - 2\mathrm{ES}_p(Z). \tag{5.10}$$

We find this equation convenient in the next subsection.

5.2 Normal risks

Here we work with the standard normal rv $Z \sim N(0,1)$ whose cdf we denote by Φ, and the p-quantile by z_p.

Corollary 5.2. *For $Z \sim N(0,1)$ and every $p \in (0,1)$, we have*

$$\mathrm{ES}_p(Z) = \frac{\Phi'(z_p)}{1-p} \tag{5.11}$$

and

$$\mathrm{TGini}_p(Z) = \frac{2(1-\Phi(\sqrt{2}z_p))}{\sqrt{\pi}(1-p)^2} - 2\mathrm{ES}_p(Z). \tag{5.12}$$

Letting $p \downarrow 0$, equation (5.12) *reduces to*

$$\mathrm{Gini}(Z) = \frac{2}{\sqrt{\pi}}. \tag{5.13}$$

Proof. The standard normal is a spherical distribution with $g(z) = \exp(-z)$. Hence $c = 1/\sqrt{2\pi}$ and $\overline{G}(z^2/2) = \Phi'(z)$. Equation (5.11), which is well known (e.g., Exercise 2.7.16 on p. 98 in Denuit et al. (2005)), follows immediately from equation (5.4).

To prove equation (5.12), we first rewrite equation (5.10) as

$$\mathrm{TGini}_p(Z) = \frac{4\mathrm{Var}(Z)}{(1-p)^2}\int_{z_p}^{\infty} f(z)f^*(z)\mathrm{d}z - 2\mathrm{ES}_p(Z) \tag{5.14}$$

with $f^*(z) = \overline{G}(z^2/2)/\mathrm{Var}(Z)$. Since $Z^* \stackrel{d}{=} Z$, we have $\mathrm{Var}(Z) = 1$. Hence, $f(z) = f^*(z) = \Phi'(z)$ and so

$$f(z)f^*(z) = \frac{1}{\sqrt{2\pi}}\Phi'\big(\sqrt{2}z\big).$$

Consequently,

$$\int_{z_p}^{\infty} f(z)f^*(z)\mathrm{d}z = \frac{1}{2\sqrt{\pi}}\big(1-\Phi(\sqrt{2}z_p)\big).$$

This establishes equation (5.12). Letting $p \downarrow 0$ in equation (5.12), we arrive at equation (5.13), which is well known (e.g., Yitzhaki and Schechtman (2013)). □

5.3 Student-t risks

Let Z be a standard Student-t rv, succinctly $Z \sim t(\theta)$, with parameter $\theta > 1/2$ and the pdf

$$f_\theta(z) = c_\theta\left(1 + \frac{z^2}{2k_\theta}\right)^{-\theta}, \quad z \in \mathbb{R}, \tag{5.15}$$

where

$$c_\theta = \frac{1}{\sqrt{2k_\theta}\,\mathrm{Beta}(1/2, \theta - 1/2)}$$

with $\mathrm{Beta}(a,b)$ denoting the classical beta function, and

$$k_\theta = \begin{cases} 1/2 & \text{when} \quad 1/2 < \theta \le 3/2, \\ \theta - 3/2 & \text{when} \quad \theta > 3/2. \end{cases}$$

Remark 5.2. The above choice of parametrization guarantees that the variance $\mathrm{Var}(Z)$, whenever it exists ($\theta > 3/2$), is equal to 1 (see McDonald (1996)). In order to obtain the standard form of the Student-t pdf from (5.15), one may choose $\theta = (1+\nu)/2$ and $k_\theta = \nu/2$, where ν denotes the degrees of freedom parameter (with the latter parametrization, $\mathrm{Var}(Z)$ is no longer 1).

We denote the cdf of $Z \sim t(\theta)$ and its p-quantile by F_θ and z_p, respectively, and note that the mean of Z is finite only if $\theta > 1$.

Corollary 5.3. *For $Z \sim t(\theta)$, $\theta > 1$ and every $p \in (0,1)$, we have*

$$\mathrm{ES}_p(Z) = \frac{c_\theta k_\theta}{(\theta-1)(1-p)} \left(1 + \frac{z_p^2}{2k_\theta}\right)^{-(\theta-1)} \tag{5.16}$$

and

$$\mathrm{TGini}_p(Z) = \frac{4c_\theta^2 k_\theta^{3/2}}{c_{2\theta-1} k_{2\theta-1}^{1/2} (\theta-1)(1-p)^2} \left(1 - F_{2\theta-1}\left(\sqrt{\frac{k_{2\theta-1}}{k_\theta}} z_p\right)\right) - 2\mathrm{ES}_p(Z). \tag{5.17}$$

Consequently,

$$\mathrm{Gini}(Z) = \frac{4c_\theta^2 k_\theta^{3/2}}{c_{2\theta-1} k_{2\theta-1}^{1/2} (\theta-1)}. \tag{5.18}$$

Proof. Student-t is a spherical distribution with the density generator $g(z) = (1 + z/k_\theta)^{-\theta}$. Hence

$$\overline{G}(z) = c_\theta \int_z^\infty \left(1 + \frac{x}{k_\theta}\right)^{-\theta} \mathrm{d}x = \frac{c_\theta k_\theta}{\theta - 1}\left(1 + \frac{z}{k_\theta}\right)^{-(\theta-1)}.$$

The variance $\mathrm{Var}(Z)$ is infinite when $1 < \theta \le 3/2$ and thus $f^*(z)$ does not exist. Nevertheless, we have

$$\begin{aligned}\overline{G}(z^2/2) f_\theta(z) &= \frac{c_\theta^2 k_\theta}{\theta - 1}\left(1 + \frac{z^2}{2k_\theta}\right)^{-(2\theta-1)} \\ &= \frac{c_\theta^2 k_\theta}{c_{2\theta-1}(\theta-1)} f_{2\theta-1}\left(\sqrt{\frac{k_{2\theta-1}}{k_\theta}} z\right)\end{aligned}$$

and then

$$\int_{z_p}^\infty \overline{G}(z^2/2) f_\theta(z) \mathrm{d}z = \frac{c_\theta^2 k_\theta}{c_{2\theta-1}(\theta-1)} \sqrt{\frac{k_\theta}{k_{2\theta-1}}} \left(1 - F_{2\theta-1}\left(\sqrt{\frac{k_{2\theta-1}}{k_\theta}} z_p\right)\right).$$

Equation (5.17) follows via equation (5.5). Consequently, when $\theta > 1$, we have the equation

$$\lim_{p \downarrow 0} \mathrm{TGini}_p(Z) = \frac{4c_\theta^2 k_\theta}{c_{2\theta-1}(\theta-1)} \sqrt{\frac{k_\theta}{k_{2\theta-1}}}$$

from which equation (5.18) follows immediately. □

Remark 5.3. Since the standard Student-t distribution converges to the standard normal distribution when $\theta \uparrow \infty$, we recover equation (5.13) by taking the limit of the right-hand side of equation (5.18) when $\theta \uparrow \infty$.

5.4 Skew-normal risks

In this section we demostrate that deriving explicit formulas for the Gini shortfall risk measure is feasible beyound the context of symmetric distributions. To this end we employ the skew-normal distributions (e.g. Azzalini (1985)). Namely, recall that rv ξ has a standard skew-normal distribution with skewness parameter $\alpha \in \mathbb{R}$, succinctly $\xi \sim SN(\alpha)$, if its pdf is given by

$$f_\xi(z) = 2\phi(z)\Phi(\alpha z),\ z \in \mathbb{R}, \tag{5.19}$$

where, as before, ϕ and Φ denote, respectively, the pdf and cdf of a standard normal rv. The following proposition is latter on employed to develop the desired expressions for $\mathrm{ES}_p(\xi)$ and $\mathrm{TGini}_p(\xi)$ with $p \in (0,\ 1)$.

For $z \in \mathbb{R}$, let $\mathbf{0}(z) = (0, 0, z)'$, and $\mathbf{0} = (0, 0)'$.

Proposition 5.4. *Let $a, b \in \mathbb{R}$ and $\sigma > 0$ be constants, and let Φ and $\Phi_n(\cdot; \Sigma)$ denote, respectively, the cdf of $Z \sim N(0, 1)$ and the cdf of an n-dimensional normal rv with zero mean vector and variance-covariance matrix Σ. Also let*

$$\Sigma_2 = \begin{pmatrix} 1 + a^2\sigma^2 & ab\sigma^2 \\ ab\sigma^2 & 1 + b^2\sigma^2 \end{pmatrix} \text{ and } \Sigma_3 = \begin{pmatrix} 1 + a^2\sigma^2 & ab\sigma^2 & -a\sigma^2 \\ ab\sigma^2 & 1 + b^2\sigma^2 & -b\sigma^2 \\ -a\sigma^2 & -b\sigma^2 & \sigma^2 \end{pmatrix}.$$

Then we have

$$\mathbb{E}[\Phi(a\sigma Z)\Phi(b\sigma Z)\mathbb{I}\{\sigma Z > z\}] = \Phi_2(\mathbf{0}; \Sigma_2) - \Phi_3(\mathbf{0}(z); \Sigma_3) \tag{5.20}$$

for every $z \in \mathbb{R}$.

Proof. Let Z^* and Z^{**} denote two independent copies of the rv Z. Then we have

$$\begin{aligned}
&\mathbb{E}[\Phi(a\sigma Z)\Phi(b\sigma Z)\mathbb{I}\{\sigma Z > z\}] \\
&= \mathbb{E}[\mathbb{P}[Z^* \le a\sigma Z,\ Z^{**} \le b\sigma Z]\mathbb{I}\{\sigma Z > z\}] \\
&= \mathbb{E}\left[\mathbb{E}[\mathbb{I}\{Z^* - a\sigma Z \le 0,\ Z^{**} - b\sigma Z \le 0\} |\ Z]\mathbb{I}\{\sigma Z > z\}\right] \\
&= \mathbb{E}\left[\mathbb{E}[\mathbb{I}\{Z^* - a\sigma Z \le 0,\ Z^{**} - b\sigma Z \le 0,\ \sigma Z > z\} |\ Z]\right] \\
&= \mathbb{P}[Z^* - a\sigma Z \le 0,\ Z^{**} - b\sigma Z \le 0,\ \sigma Z > z] \\
&= \mathbb{P}[Z^* - a\sigma Z \le 0,\ Z^{**} - b\sigma Z \le 0] - \mathbb{P}[Z^* - a\sigma Z \le 0,\ Z^{**} - b\sigma Z \le 0,\ \sigma Z \le z],
\end{aligned}$$

and the desired formula follows because the class of normal rv's is closed under affine transformations. □

Further consider the following matrices

$$\Sigma_{12} = \begin{pmatrix} 1+\alpha^2/2 & \alpha^2/2 \\ \alpha^2/2 & 1+\alpha^2/2 \end{pmatrix}, \ \Sigma_{13} = \begin{pmatrix} 1+\alpha^2/2 & \alpha^2/2 & -\alpha/2 \\ \alpha^2/2 & 1+\alpha^2/2 & -\alpha/2 \\ -\alpha/2 & -\alpha/2 & 1/2 \end{pmatrix},$$

$$\Sigma_{22} = \begin{pmatrix} 1+\alpha^2 & \alpha\sqrt{1+\alpha^2} \\ \alpha\sqrt{1+\alpha^2} & 2+\alpha^2 \end{pmatrix}, \text{ and } \Sigma_{23} = \begin{pmatrix} 1+\alpha^2 & \alpha\sqrt{1+\alpha^2} & -\alpha \\ \alpha\sqrt{1+\alpha^2} & 2+\alpha^2 & -\sqrt{1+\alpha^2} \\ -\alpha & -\sqrt{1+\alpha^2} & 1 \end{pmatrix}.$$

Corollary 5.5. *For $\xi \sim SN(\alpha)$, its p-quantile ξ_p with $p \in (0,1)$, we have*

$$\mathrm{ES}_p(\xi) = \frac{2}{1-p}\left(\Phi(\alpha\xi_p)\phi(\xi_p) + \frac{\alpha}{\sqrt{2\pi(1+\alpha^2)}}\left(1 - \Phi\left(\xi_p\sqrt{1+\alpha^2}\right)\right)\right) \tag{5.21}$$

and

$$\begin{aligned}\mathrm{TGini}_p(\xi) = \frac{8}{(1-p)^2}&\left(F_\xi(\xi_p)\Phi(\alpha\xi_p)\phi(\xi_p) + \frac{\Phi_2(\mathbf{0};\Sigma_{12}) - \Phi_3(\mathbf{0}(\xi_p);\Sigma_{13})}{\sqrt{\pi}}\right.\\ &\left.+\frac{1 - F_\xi(\xi_p)\Phi\left(\xi_p\sqrt{1+\alpha^2}\right) - 2\Phi_2(\mathbf{0};\Sigma_{22}) + 2\Phi_3(\mathbf{0}(\xi_p);\Sigma_{23})}{\sqrt{2\pi(1+\alpha^2)}}\right)\\ &- \frac{2(1+p)}{1-p}\mathrm{ES}_p(\xi).\end{aligned} \tag{5.22}$$

Consequently,

$$\mathrm{Gini}(\xi) = 8\left(\frac{\Phi_2(\mathbf{0};\Sigma_{12})}{\sqrt{\pi}} + \frac{1 - 2\Phi_2(\mathbf{0};\Sigma_{22})}{\sqrt{2\pi(1+\alpha^2)}}\right) - 2\alpha\sqrt{\frac{2}{\pi(1+\alpha^2}}. \tag{5.23}$$

Proof. To obtain (5.21) we integrate by parts and have

$$\begin{aligned}\mathrm{ES}_p(\xi) = \frac{1}{1-p}\int_{\xi_p}^{\infty} z f_\xi(z)\mathrm{d}z &= \frac{2}{1-p}\int_{\xi_p}^{\infty} z\phi(z)\Phi(\alpha z)\mathrm{d}z\\ &= \frac{2}{1-p}\left(\Phi(\alpha\xi_p)\phi(\xi_p) + \alpha\int_{\xi_p}^{\infty}\phi(\alpha z)\phi(z)\mathrm{d}z\right)\\ &= \frac{2}{1-p}\left(\Phi(\alpha\xi_p)\phi(\xi_p) + \frac{\alpha}{\sqrt{2\pi(1+\alpha^2)}}\mathbb{E}\left[\mathbb{I}\{Z_0 > \xi_p\}\right]\right),\end{aligned}$$

where Z_0 is a normal rv with zero mean and variance $1/(1+\alpha^2)$. Formula (5.21) then follows by evoking Proposition 5.4 with $a = 0$, $b = 0$ and $\sigma = 1/\sqrt{1+\alpha^2}$. We note in passing that when $\alpha = 0$, that is the skewness parameter is equal to zero and so $\xi \sim N(0,1)$, then (5.21) reduces to the expression derived in Panjer (2001). In addition, when $p \downarrow 0$, we readily obtain that

$$\lim_{p\downarrow 0} \mathrm{ES}_p(\xi) = \mathbb{E}[\xi] = \alpha\sqrt{\frac{2}{\pi(1+\alpha^2)}},$$

which confirms the findings of Azzalini (1985).

To further compute the TGini_p functional, we recall (3.8), and so our main target is

$$\mathbb{E}[\xi F_\xi(\xi)\mathbb{I}\{\xi > \xi_p\}] = \int_{\xi_p}^{\infty} zF_\xi(z)f_\xi(z)\mathrm{d}z = 2\int_{\xi_p}^{\infty} zF_\xi(z)\phi(z)\Phi(\alpha z)\mathrm{d}z,$$

which after integration by parts reduces to

$$\begin{aligned}&\mathbb{E}[\xi F_\xi(\xi)\mathbb{I}\{\xi > \xi_p\}]\\ &= 2\left(F_\xi(\xi_p)\Phi(\alpha\xi_p)\phi(\xi_p) + \int_{\xi_p}^{\infty} f_\xi(z)\Phi(\alpha z)\phi(z)\mathrm{d}z + \alpha\int_{\xi_p}^{\infty} F_\xi(z)\phi(\alpha z)\phi(z)\mathrm{d}z\right).\end{aligned}$$

In order to compute the two integrals in the last expression we employ Proposition 5.4. Specifically, for Z_1 being a normal rv with zero mean and variance $1/2$, we use the following change of measure

$$\begin{aligned}\mathbb{E}[\Phi(\alpha\xi)\phi(\xi)\mathbb{I}\{\xi > \xi_p\}] &= \frac{1}{\sqrt{\pi}}\mathbb{E}\left[\Phi(\alpha Z_1)^2\mathbb{I}\{Z_1 > \xi_p\}\right]\\ &= \frac{1}{\sqrt{\pi}}\mathbb{E}\left[\Phi\left(\alpha\frac{1}{\sqrt{2}}Z\right)^2\mathbb{I}\left\{\frac{1}{\sqrt{2}}Z > \xi_p\right\}\right]\end{aligned}$$

and further evoke Proposition 5.4 with $a = b = \alpha$ and $\sigma = 1/\sqrt{2}$ to obtain

$$\begin{aligned}&\sqrt{\pi}\int_{\xi_p}^{\infty} f_\xi(z)\Phi(\alpha z)\phi(z)\mathrm{d}z\\ &= \Phi_2\left(\mathbf{0};\begin{pmatrix}1+\alpha^2/2 & \alpha^2/2\\ \alpha^2/2 & 1+\alpha^2/2\end{pmatrix}\right) - \Phi_3\left(\mathbf{0}(\xi_p);\begin{pmatrix}1+\alpha^2/2 & \alpha^2/2 & -\alpha/2\\ \alpha^2/2 & 1+\alpha^2/2 & -\alpha/2\\ -\alpha/2 & -\alpha/2 & 1/2\end{pmatrix}\right).\end{aligned}$$

In a similar fashion, we integrate by parts and have

$$\begin{aligned}&\sqrt{2\pi(1+\alpha^2)}\int_{\xi_p}^{\infty} F_\xi(z)\phi(\lambda z)\phi(z)\mathrm{d}z\\ &= 1 - F_\xi(\xi_p)\Phi\left(\xi_p\sqrt{1+\alpha^2}\right) - 2\mathbb{E}\left[\Phi(\alpha Z)\Phi(Z\sqrt{1+\alpha^2})\mathbb{I}\{Z > \xi_p\}\right],\end{aligned}$$

which after setting $a=\alpha, b=\sqrt{1+\alpha^2}$ and $\sigma=1$ in Proposition 5.4 implies

$$\begin{aligned}
&\sqrt{2\pi(1+\alpha^2)}\int_{\xi_p}^{\infty} F_\xi(z)\phi(\lambda z)\phi(z)\mathrm{d}z \\
&= 1-F_\xi(\xi_p)\Phi\left(\xi_p\sqrt{1+\alpha^2}\right)-2\Phi_2\left(\mathbf{0};\begin{pmatrix} 1+\alpha^2 & \alpha\sqrt{1+\alpha^2} \\ \alpha\sqrt{1+\alpha^2} & 2+\alpha^2\end{pmatrix}\right) \\
&\quad +2\Phi_3\left(\mathbf{0}(\xi_p);\begin{pmatrix} 1+\alpha^2 & \alpha\sqrt{1+\alpha^2} & -\alpha \\ \alpha\sqrt{1+\alpha^2} & 2+\alpha^2 & -\sqrt{1+\alpha^2} \\ -\alpha & -\sqrt{1+\alpha^2} & 1\end{pmatrix}\right).
\end{aligned}$$

This proves (5.22), from which formula (5.23) follows as a simple limiting case. □

5.5 Skew-t risks

A natural generalization of the skew-normal distribution that encompasses heavy-tailed risks is the skew-t distribution. Let f_ν and F_ν denote, respectively, the pdf and cdf of the standard Student-t rv with the degrees of freedom parameter $\nu>0$; that is, f_ν follows from (5.15) by setting $\theta=(\nu+1)/2$ and $k_\theta=\nu/2$ as

$$f_\nu(z)=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\nu\pi}}\left(1+\frac{z^2}{\nu}\right)^{-(\nu+1)/2},\ z\in\mathbb{R}.$$

Then the standard skew-t rv, succinctly $\xi\sim St(\alpha,\nu)$ with skewness $\alpha\in\mathbb{R}$ and degrees of freedom $\nu>0$, has the pdf (see e.g. Azzalini and Capitanio (2003))

$$f_\xi(z)=2f_\nu(z)F_\nu\left(\alpha z\sqrt{\frac{\nu+1}{\nu+z^2}};\nu+1\right),\ z\in\mathbb{R}.$$

We further provide a number of figures depicting the ES and TGini risk measures in the context of the skew-t risk rv's. We note in passing that even for the skew-normal distributions, that are the limiting case of the skew-t distributions for $\nu\uparrow\infty$, the formulas for ES and TGini are rather involved (Section 5.4). In the case of the skew-t distributions, the task of developing analytical expressions for ES and TGini is noticeably more cumbersome. For this reason, in this section we have opted for the Monte-Carlo simulation approach, and to this end we have embarked on equation (3.6), as well as on the stochastic representation of the skew-t rv's as scale mixtures of the skew-normal rv's (Azzalini and Capitanio (2003)).

In Figures 1 and 2, values of ES_p, TGini_p and $\mathrm{GS}_p^{1/2}$ for skew-t risks are reported for several choices of α and ν. All calculations are carried out via simulation of sample size 10^6. It is clear from Figures 1 and 2 that, compared to the expected shortfall, the

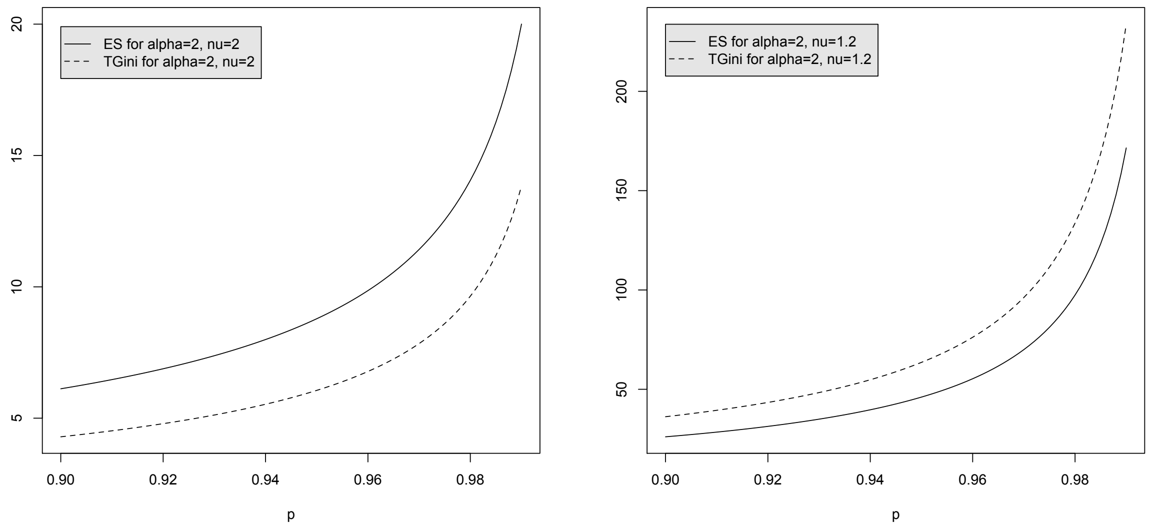

Figure 1: ES_p and TGini_p, $p \in [0.9, 0.99]$ for skew-t risks with $\alpha = 2$ and $\nu = 2$ (left) and $\alpha = 2$ and $\nu = 1.2$ (right)

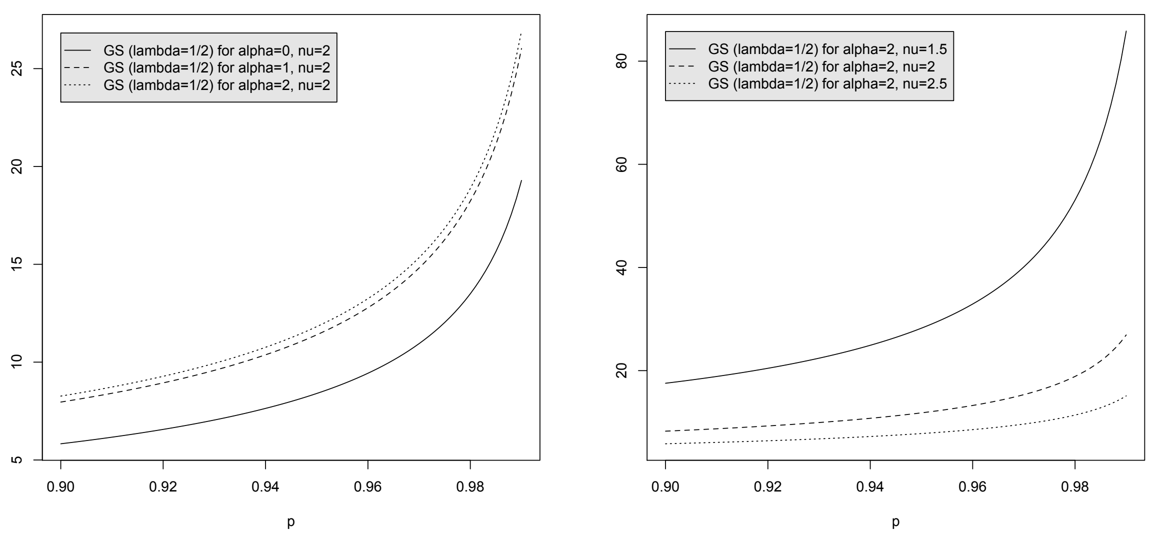

Figure 2: $\mathrm{GS}_p^{1/2}$, $p \in [0.9, 0.99]$ for skew-t risks with $\alpha = 0, 1, 2$, $\nu = 2$ (left) and $\alpha = 2$, $\nu = 1.5, 2, 2.5$ (right)

TGini functional (and hence the Gini shortfall) is more sensitive to the degrees of freedom parameter ν, which represents the heaviness of tail risk. In particular, if ν is close to 1, the TGini functional is larger than the corresponding expected shortfall; it is the other way around for larger values of ν (e.g. $\nu \geq 2$).

6 Gini shortfall allocation

6.1 Intuitive definition

Driven by the recent regulatory frameworks (e.g., Cruz (2009), Sandström (2010), Cannata and Quagliariello (2011)), here we introduce a capital allocation counterpart to the Gini shortfall GS_p^λ.

For any portfolio $\mathbf{X} = (X_1, \ldots, X_n)' \in \mathcal{X}^n$ and its aggregate risk $S = \sum_{k=1}^n X_i$, the aim is to allocate the total capital $\mathrm{GS}_p^\lambda(S)$ to n constituents corresponding to $X_1, \ldots, X_n$. A natural idea for such allocation hinges on appropriate extensions of the earlier defined functionals $X \mapsto \mathrm{ES}_p(X)$ and $X \mapsto \mathrm{TGini}_p(X)$. For this, we first introduce additional notation. Namely, let U_S be the distributional transform of S (Proposition 1.3 of Rüschendorf (2013)) defined by

$$U_S = F_S(S-) + V(F_S(S) - F_S(S-)), \tag{6.1}$$

where V is a uniform on $[0,1]$ rv independent of $X_1, \ldots, X_n$. This implies the uniform on $[0,1]$ distribution for U_S and ensures the equation $F_S^{-1}(U_S) = S$ almost surely. The aforementioned extensions of the expected shortfall and the tail-Gini are, for $k = 1, \ldots, n$, then defined as follows:

$$\mathrm{ES}_p(X_k, S) = \mathbb{E}[X_k \mid U_S > p] \tag{6.2}$$

and

$$\mathrm{TGini}_p(X_k, S) = \frac{4}{1-p}\mathrm{Cov}[X_k, U_S \mid U_S > p]. \tag{6.3}$$

Functional (6.2) has recently been employed (e.g. Acharya et. al. (2012)) to measure systemic risk (SRISK). In a more general context of weighted capital allocations, functional (6.2) has been explored in detail by Furman and Zitikis (2008, 2009).

Note that when $\mathbb{P}(S = s_p) = 0$ with $s_p = F_S^{-1}(p)$, which is the case for elliptical portfolios to be considered in Section 6.2 below, definitions (6.2) and (6.3) simplify because we do not need to involve the distributional transform U_S. Namely, we have

$$\mathrm{ES}_p(X_k, S) = \mathbb{E}[X_k \mid S > s_p] \tag{6.4}$$

and

$$\mathrm{TGini}_p(X_k, S) = \frac{4}{1-p}\mathrm{Cov}[X_k, F_S(S) \mid S > s_p]. \tag{6.5}$$

Mimicking equation (4.1), we next define the Gini shortfall allocation by the equation

$$\mathrm{GS}_p^\lambda(X_k, S) = \mathrm{ES}_p(X_k, S) + \lambda\ \mathrm{TGini}_p(X_k, S), \tag{6.6}$$

where $p \in [0, 1)$ is the prudence level, and $\lambda \geq 0$ is the loading parameter. In Section 6.2 below, we illustrate this Gini shortfall allocation by deriving closed-form expressions for the elliptical portfolio of risks. In this case, S has a continuous distribution and we can thus rely on formulas (6.4) and (6.5). Note the obvious though important equation $\sum_{k=1}^n \mathrm{GS}_p^\lambda(X_k, S) = \mathrm{GS}_p^\lambda(S, S)$, with the right-hand side equal to the Gini shortfall $\mathrm{GS}_p^\lambda(S)$ given by equation (4.1) with the aggregate risk S in the role of X.

Remark 6.1. Capital allocation rules (6.2)-(6.6) coincide with the corresponding Euler allocation principles (see Section 8.5 of McNeil et al. (2015)) when some regularity of the joint distribution of $(X_1, \ldots, X_n, S)$ is assumed; see Proposition 1 of Tsanakas and Millossovich (2016).

6.2 Aggregate elliptical risks

Let $\mathbf{X} = (X_1, \ldots, X_n)' \in \mathcal{X}^n$ be a portfolio of elliptical risks with the vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ of finite expectations, a positive-definite symmetric matrix B, and the pdf

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{c_n}{\sqrt{|B|}} g_n\left(\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'B^{-1}(\mathbf{x}-\boldsymbol{\mu})\right),\ \mathbf{x} \in \mathbb{R}^n,$$

where g_n is an n-dimensional density generator, and c_n is a normalizing constant. Succinctly, we write $\mathbf{X} \sim E_n(\boldsymbol{\mu}, B, g_n)$. For the aggregate risk $S = X_1 + \cdots + X_n$, let $\mu_S = \boldsymbol{\mu}'\mathbf{1}$ and $\beta_S^2 = \mathbf{1}'B\mathbf{1}$, where $\mathbf{1} = (1, \ldots, 1)'$ is the n-dimensional vector of 1's. The next theorem provides formulas for calculating the Gini shortfall $\mathrm{GS}_p^\lambda(S)$ of the aggregate risk S.

Let g denote the univariate density generator corresponding to g_n (see Fang et al. (1987) for details). Recall that $Z \sim S(g)$ is a spherical rv, z_p denotes its p-quantile, and

$$\overline{G}(y) = c\int_y^\infty g(x)\mathrm{d}x$$

is the tail generator, which is well defined because we assume that $\mathbb{E}[Z] < \infty$.

Theorem 6.1. *When $\mathbf{X} \sim E_n(\boldsymbol{\mu}, B, g_n)$, then for every $p \in (0, 1)$ we have*

$$\mathrm{ES}_p(S) = \mu_S + \frac{\overline{G}(z_p^2/2)}{1-p}\beta_S \tag{6.7}$$

and

$$\mathrm{TGini}_p(S) = \left(\frac{4}{1-p} \mathbb{E}\left[\overline{G}(Z^2/2) \mid Z > z_p \right] - 2\mathrm{ES}_p(Z) \right) \beta_S. \tag{6.8}$$

Letting $p \downarrow 0$ *in equation* (6.8)*, we obtain*

$$\mathrm{Gini}(S) = 4\mathbb{E}\left[\overline{G}(Z^2/2) \right] \beta_S. \tag{6.9}$$

Proof. Recall (e.g., Fang et al. (1987)) that $S \sim E_1(\mu_S, \beta_S, g)$. Hence, equations (6.7) and (6.9) follow from Theorem 5.1 as follows

$$\mathrm{ES}_p(S) = \mathrm{ES}_p(\mu_S + \beta_S Z) = \mu_S + \beta_S\ \mathrm{ES}_p[Z]$$

and

$$\mathrm{TGini}_p(S) = \mathrm{TGini}_p(\mu_S + \beta_S Z) = \beta_S\ \mathrm{TGini}_p(Z).$$

This completes the proof of Theorem 6.1. □

Remark 6.2. Theorem 6.1 implies that, similarly to VaR_q, $q \in [1/2, 1)$, TGini_p, $p \in (0, 1)$ is sub-additive for jointly elliptical risks; note that VaR_q and TGini_p are not sub-additive in general (Proposition 3.3). By Theorem 8.28 of McNeil et al. (2015), all positively homogeneous, translation-invariant and law-invariant risk measures no less than the mean are sub-additive for jointly elliptical risks; this applies to GS_p^λ for all $\lambda \geq 0$ and $p \in (0, 1)$.

We next derive formulas for the Gini shortfall allocation, and our task mainly hinges on deriving expressions for $\mathrm{ES}_p(X_k, S)$ and $\mathrm{TGini}_p(X_k, S)$. To this end, assume that the aforementioned matrix B has diagonal entries β_k^2 and off-diagonal entries $\beta_{k,l} = \beta_{l,k}$, $k, l = 1, \ldots, n$, and recall the following well-known regression formula that holds for $\mathbf{X} \sim E_n(\boldsymbol{\mu}, B, g_n)$ (e.g., Fang et al. (1987))

$$\mathbb{E}[X_k \mid S = s] = \mu_k + \frac{\beta_{k,S}}{\beta_S^2} \big(s - \mu_S \big), \quad s \in \mathbb{R}, \tag{6.10}$$

where $\beta_{k,S} = \beta_{k,1} + \cdots + \beta_{k,n}$.

Theorem 6.2. *Let* $\mathbf{X} \sim E_n(\boldsymbol{\mu}, B, g_n)$*. For every* $p \in (0, 1)$ *and* $k = 1, \ldots, n$*, we have*

$$\mathrm{ES}_p(X_k, S) = \mu_k + \frac{\beta_{k,S}}{\beta_S^2} \big(\mathrm{ES}_p[S] - \mu_S \big), \tag{6.11}$$

and

$$\mathrm{TGini}_p(X_k, S) = \frac{\beta_{k,S}}{\beta_S} \left(\frac{4}{1-p} \mathbb{E}\left[\overline{G}(Z^2/2) \mid Z > z_p \right] - 2\mathrm{ES}_p(Z) \right). \tag{6.12}$$

Letting $p \downarrow 0$ *in equation* (6.12)*, we obtain*

$$\mathrm{Gini}(X_k, S) = 4\mathrm{Cov}[X_k, F_S(S)] = 4 \frac{\beta_{k,S}}{\beta_S} \mathbb{E}\left[\overline{G}(Z^2/2) \right]. \tag{6.13}$$

Proof. Equation (6.11) follows immediately from equation (6.10), with a formula for $\mathrm{ES}_p(S)$ given in Theorem 6.1. Hence, we only need to verify formula (6.12). We start with the equations

$$\begin{aligned}\mathrm{TGini}_p(X_k,S) &= \frac{4}{1-p}\mathbb{E}[(X_k-\mathrm{ES}_p(X_k,S))F_S(S)\mid S>s_p]\\ &= \frac{4}{1-p}\mathbb{E}[(X_k-\mathbb{E}[X_k])F_S(S)\mid S>s_p]\\ &\quad -\frac{4}{1-p}(\mathrm{ES}_p(X_k,S)-\mathbb{E}[X_k])\mathbb{E}[F_S(S)\mid S>s_p]. \end{aligned}\tag{6.14}$$

Equation (6.10) implies

$$\begin{aligned}&\mathbb{E}[(X_k-\mathrm{ES}_p(X_k,S))F_S(S)\mid S>s_p]\\ &=\frac{\beta_{k,S}}{\beta_S^2}\mathbb{E}[(S-\mathbb{E}[S])F_S(S)\mid S>s_p]\\ &=\frac{\beta_{k,S}}{\beta_S^2}\mathbb{E}[(S-\mathrm{ES}_p(S))F_S(S)\mid S>s_p]+\frac{\beta_{k,S}}{\beta_S^2}(\mathrm{ES}_p(S)-\mathbb{E}[S])\mathbb{E}[F_S(S)\mid S>s_p]\\ &=\frac{\beta_{k,S}}{\beta_S^2}\mathrm{Cov}[S,F_S(S)\mid S>s_p]+\frac{\beta_{k,S}}{\beta_S^2}(\mathrm{ES}_p(S)-\mathbb{E}[S])\frac{1+p}{2}. \end{aligned}\tag{6.15}$$

With the help of Theorem 6.1 for calculating the quantities on the right-hand side of equation (6.15), we obtain

$$\begin{aligned}&\frac{4}{1-p}\mathbb{E}[(X_k-\mathrm{ES}_p(X_k,S))F_S(S)\mid S>s_p]\\ &\qquad=\frac{\beta_{k,S}}{\beta_S}\left(\frac{4}{1-p}\mathbb{E}\left[\overline{G}(Z^2/2)\mid Z>z_p\right]-2\mathrm{ES}_p(Z)\right)+\frac{2(1+p)}{(1-p)^2}\frac{\beta_{k,S}}{\beta_S}\overline{G}(z_p^2/2). \end{aligned}\tag{6.16}$$

Upon recalling equation (6.11), we have

$$\begin{aligned}\frac{4}{1-p}(\mathrm{ES}_p(X_k,S)-\mathbb{E}[X_k])\mathbb{E}[F_S(S)\mid S>s_p] &= \frac{2(1+p)}{1-p}(\mathrm{ES}_p(X_k,S)-\mathbb{E}[X_k])\\ &=\frac{2(1+p)}{1-p}\frac{\beta_{k,S}}{\beta_S^2}\big(\mathrm{ES}_p[S]-\mathbb{E}[S]\big)\\ &=\frac{2(1+p)}{(1-p)^2}\frac{\beta_{k,S}}{\beta_S}\overline{G}(z_p^2/2). \end{aligned}\tag{6.17}$$

Using formulas (6.16) and (6.17) on the right-hand side of equation (6.14), we arrive at equation (6.12). This finishes the proof of Theorem 6.2. □

7 An application

We consider a bancassurance company with ten business lines, as described in Panjer (2001). The rv's of interest represent the present values of the amounts that are required to guarantee solvency over a fixed time horizon with a high confidence level. We find it reasonable to assume that the joint distribution is the Student-t distribution (Section 5.3 for details) with the mean vector

$$\boldsymbol{\mu} = (25.69, 37.84, 0.85, 12.70, 0.15, 24.05, 14.41, 4.49, 4.39, 9.56)'$$

and the positive definite matrix

$$B = \begin{pmatrix} 7.24 & 0 & 0.07 & -0.07 & 0.28 & -2.71 & -0.51 & 0.28 & 0.23 & -0.21 \\ 0 & 20.16 & 0.05 & 1.60 & 0.05 & 1.39 & 1.14 & -0.91 & -0.81 & -1.74 \\ 0.07 & 0.05 & 0.04 & 0.00 & -0.01 & 0.08 & 0.01 & -0.02 & -0.02 & -0.07 \\ -0.07 & 1.60 & 0.00 & 1.74 & 0.17 & 0.26 & 0.19 & -0.14 & 0.18 & -0.79 \\ 0.28 & 0.05 & -0.01 & 0.17 & 0.32 & -0.24 & 0.01 & -0.02 & 0.08 & -0.01 \\ -2.71 & 1.39 & 0.08 & 0.26 & -0.24 & 14.98 & 0.43 & -0.33 & -1.89 & -1.60 \\ -0.51 & 1.14 & 0.01 & 0.19 & 0.01 & 0.43 & 2.53 & -0.38 & 0.13 & 0.58 \\ 0.28 & -0.91 & -0.02 & -0.14 & -0.02 & -0.33 & -0.38 & 0.92 & -0.16 & -0.40 \\ 0.23 & -0.81 & -0.02 & 0.18 & 0.08 & -1.89 & 0.13 & -0.16 & 1.12 & 0.58 \\ -0.21 & -1.74 & -0.07 & -0.79 & -0.01 & -1.60 & 0.58 & -0.40 & 0.58 & 6.71 \end{pmatrix}.$$

Because of the special parametrization of the Student-t distribution adopted in Section 5.3, whenever the variance is finite, it is equal to 1, and thus the matrix B is the variance-covariance matrix for all $\theta > 3/2$. Letting $\theta \uparrow \infty$ yields the normal distribution. The Student-t distribution inherits properties of the class of elliptical distributions, and thus when $\mathbf{X} \sim t_n(\boldsymbol{\mu}, B, q)$, then X_k and S jointly follow the two-dimensional Student-t distribution $t_2(\boldsymbol{\mu}_{k,S}, B_{k,S}, q)$ with $\boldsymbol{\mu}_{k,S} = A'_k\boldsymbol{\mu}$ and $B_{k,S} = A'_k B A_k$, where

$$A_k = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \overbrace{1}^{k\text{-th}} & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}'.$$

This way we obtain the vector

$$\beta_{k,S} = (4.52, 20.84, 0.13, 3.20, 0.61, 10.41, 4.17, -1.16, -0.50, 3.14)'$$

whose entries are the off-diagonal elements of the positive-definite matrix $B_{k,S}$. When $B_{k,S}$ is a covariance matrix, then the Pearson correlations of the risk due to the k-th business line and the aggregate risk S are

$$\rho_{k,S} = (0.25, 0.69, 0.09, 0.36, 0.16, 0.40, 0.39, -0.18, -0.07, 0.18)'.$$

In what follows we apply the earlier introduced Gini-type risk measures and allocation rules in three contexts, which make up Sections 7.1–7.3 that we now briefly overview:

- In Section 7.1 we discuss pricing the portfolio constituents when the risks are considered stand alone and when they are pulled. We find that the Gini shortfall encourages diversification at a rate that is higher than that of the expected shortfall.

- In Section 7.2 we calculate the allocated economic capitals in the context of our portfolio of ten risks. We see that the stand alone risks are significantly more expensive than the combined ones and, also, that two risks – which are #8 and #9 – are negatively correlated with the aggregate portfolio risk and thus require negative economic capital.

- In Section 7.3, we evaluate the risk margins required for the aggregate risk of the portfolio. Often in practice, insurance companies estimate the risk margins using the value-at-risk at the prudence level $p = 0.75$. We discover that the risk margins derived from this rule are significantly underestimated in particular when the underlying risks have heavier tails than those of the normal distribution.

7.1 Pricing

We already noted in Section 1 that the tail-standard-deviation/standard-deviation shortfall risk measures of Furman and Landsman (2006a) cannot price risks with infinite second moments. Hence, we use the Gini shortfall. In Table 1 we report our findings for the aforementioned portfolio of ten risks. Note that the Gini shortfall is more supportive when it comes to diversification than the expected shortfall. Also, the expected shortfall seems to be less sensitive to the tail risk as it finds the Student-t risk with $\theta = 2$ less expensive than the normally distributed risk, although the tail of the former risk is heavier. The prices obtained with the help of the standard-deviation shortfall

$$\mathrm{SDS}_p^{\lambda}(X) = \mathrm{ES}_p(X) + \lambda \sqrt{\mathbb{E}[(X - \mathrm{ES}_p(X))^2 \mid X > x_p]}$$

and Gini shortfall risk measures are very close, and thus the latter risk measure seems to provide a good substitute for the former one in situations when the second moment is infinite, e.g., for Student-t risks with $\theta = 1.5$. We note in passing that the ES_p risk measure was used for pricing insurance risks in Furman and Landsman (2006b).

7.2 Economic capital allocation

Once the aggregate economic capital has been determined, it is usually in the interest of upper management to learn how this economic capital is allocated to different sources of

		Lines of business							Total	DIV
		1	2	3	$\cdots$	8	9	10		
$\theta = 1.5$	SDS_0^λ	NaN	NaN	NaN	$\cdots$	NaN	NaN	NaN	NaN	NaN
	ES_p	30.35	45.62	1.21	$\cdots$	6.15	6.23	14.05	145.78	0.15
	SDS_p^λ	NaN	NaN	NaN	$\cdots$	NaN	NaN	NaN	NaN	NaN
	GS_p^λ	34.25	52.12	1.52	$\cdots$	7.54	7.76	17.80	155.53	0.26
$\theta = 2$	SDS_0^λ	28.38	42.33	1.06	$\cdots$	5.45	5.45	12.15	140.86	0.09
	ES_p	28.56	42.62	1.07	$\cdots$	5.51	5.52	12.32	141.30	0.10
	SDS_p^λ	30.97	46.65	1.26	$\cdots$	6.37	6.47	14.64	147.33	0.17
	GS_p^λ	30.40	45.70	1.22	$\cdots$	6.17	6.25	14.10	145.91	0.15
$\theta = \infty$	SDS_0^λ	28.38	42.33	1.06	$\cdots$	5.45	5.45	12.15	140.86	0.09
	ES_p	29.11	43.54	1.12	$\cdots$	5.71	5.74	12.85	142.68	0.11
	SDS_p^λ	30.44	45.76	1.22	$\cdots$	6.18	6.26	14.13	146.00	0.15
	GS_p^λ	30.53	45.92	1.23	$\cdots$	6.22	6.30	14.22	146.24	0.16

Table 1: Risk measures for the Student-t risks with varying parameter θ, as well as $p = 0.75$, $\lambda = 1$, and the diversification per unit of risk (DIV).

riskiness, such as business lines in a financial enterprise. Below we present an allocation that is based on the weighted insurance pricing model (WIPM) introduced by Furman and Zitikis (2009). We note, that this allocation is akin to the Capital Asset Pricing Model's (CAPM) "beta", but unlike the CAPM, the WIPM does not require the finiteness of the second moment. In the context of Student-t portfolios, the WIPM's allocation is given by

$$w_{k,S} = \frac{\beta_{k,S}}{\beta_S^2},$$

which is precisely the CAPM's beta when $\beta_{k,S} = \mathrm{Cov}[X_k, S]$ and $\beta_S^2 = \mathrm{Var}(S)$, which happens when $\theta > 1.5$. Table 2 reports the values of $w_{k,S}$ corresponding to the earlier introduced portfolio of ten risks. Note that the standard-deviation shortfall cannot be used because the second moment is infinite when $\theta = 1.5$, which we have set in the table.

7.3 Risk margins

As mentioned earlier, the quantile approach frequently sets the risk margin to the value-at-risk at the level $p = 0.75$. Table 3 shows that this may become particularly insufficient for distributions with tails that are heavier than the tail of the normal distribution. More specifically, the risk margin that results from $\mathrm{VaR}_{0.75}$ is 18% of the risk margin based on Gini shortfall for Student-t with $\theta = 1.5$ risks, 25% of the risk margin associated with

Cost of capital k	1	2	3	4	5	6	7	8	9	10
$w_{k,S}$	0.10	0.46	0.01	0.07	0.01	0.23	0.09	-0.03	-0.01	0.07
$\mathrm{ES}_p(X_k, S) - \mathbb{E}[X_k]$	6.69	30.78	0.67	4.68	0.67	15.39	6.02	-2.01	-0.67	4.68
$\mathrm{ES}_p(X_k) - \mathbb{E}[X_k]$	26.77	44.67	2.09	13.13	5.67	38.51	15.82	9.55	10.55	25.70
$\mathrm{GS}_p^\lambda(X_k, S) - \mathbb{E}[X_k]$	11.19	51.47	1.12	7.83	1.12	25.74	10.07	-3.36	-1.12	7.83
$\mathrm{GS}_p^\lambda(X_k) - \mathbb{E}[X_k]$	44.75	74.70	3.49	21.96	9.48	64.38	26.45	15.97	17.63	43.09

Table 2: Economic capital allocations according to the expected shortfall and Gini shortfall allocation rules alongside the corresponding stand-alone economic capitals when $\theta = 1.5$, $p = 0.99$ and $\lambda = 1$.

standard deviation shortfall for Student-t with $\theta = 2$ risks, and 39% of the risk margin corresponding to Gini shortfall for normally distributed risks.

Risk margin per unit of risk	$\theta = 1.5$	$\theta = 2$	$\theta = \infty$
$\mathrm{SDS}_0^\lambda(S)/\mathbb{E}[S] - 1$	NaN	0.0502	0.0502
$\mathrm{VaR}_p(S)/\mathbb{E}[S] - 1$	0.0290	0.0221	0.0338
$\mathrm{ES}_p(S)/\mathbb{E}[S] - 1$	0.0869	0.0535	0.0637
$\mathrm{SDS}_p^\lambda(S)/\mathbb{E}[S] - 1$	NaN	0.0984	0.0885
$\mathrm{GS}_p^\lambda(S)/\mathbb{E}[S] - 1$	0.1595	0.0878	0.0903

Table 3: Risk margins per unit of risk for several risk measures in the case of Student-t risks with varying parameter θ, the prudence parameter $p = 0.75$, and $\lambda = 1$.

Acknowledgements

The authors thank the Associate Editor and an anonymous referee for various constructive comments on an early version of the paper. The authors gratefully acknowledge the support of their research by the Natural Sciences and Engineering Research Council (NSERC) of Canada. EF and RZ also acknowledge the support of their research by the Casualty Actuarial Society (CAS).

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, **26**, 1505–1518.

Acharya, V., Engle, R. and Richardson, M. (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review*, **102**, 59–64.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228.

Azzalini, A. (1985). A class of distributions which includes normal ones. *Scandinavian Journal of Statistics*, **12**, 171–178.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution *Journal of the Royal Statistical Society: Series B*, **65**, 367–389.

Cannata, F. and Quagliariello, M. (2011). *Basel III and Beyond.* Risk Books, London.

Ceriani, L. and Verme, P. (2012). The origins of the Gini index: extracts from Variabilità e Mutabilità (1912) by Corrado Gini. *Journal of Economic Inequality*, **10**, 421–443.

Cruz, M. (2009). *The Solvency II Handbook.* Risk Books, London.

Dana, R.-A. (2005). A representation result for concave Schur-concave functions. *Mathematical Finance*, **15**, 613–634.

Delbaen, F. (2012). *Monetary Utility Functions.* Osaka University Press, Osaka.

Denneberg, D. (1990). Premium calculation: why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin: Journal of the International Actuarial Association*, **20**, 181–190.

Denneberg, D. (1994). *Non-additive Measure and Integral.* Kluwer, Dordrecht.

Denuit, M., Dhaene, J., Goovaerts, M.J. and Kaas, R. (2005). *Actuarial Theory for Dependent Risk: Measures, Orders and Models.* Wiley, New York.

De Waegenaere, A. and Wakker, P.P. (2001). Nonmonotonic Choquet integrals. *Journal of Mathematical Economics*, **36**, 45–60.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. and Vyncke, D. (2002). The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, **31**, 3–33.

Dobrushin, R.L. (1970). Describing a system of rv's by conditional distributions. *Theory of Probability and its Applications*, **15**, 458–486.

Dowd, K., Cotter, J. and Sorwar, G. (2008). Spectral risk measures: properties and limitations. *Journal of Financial Services Research*, **34**(1), 61–75.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A. (2014). An academic response to Basel 3.5. *Risks*, **2**, 25-48.

Embrechts, P., Wang, B. and Wang, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, **19**, 763–790.

Emmer, S., Kratz, M. and Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, **18**, 31–60.

Fang, K., Kotz, S. and Ng, K.W. (1987). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.

Föllmer, H. and Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**, 429–447.

Föllmer, H. and Schied, A. (2011). *Stochastic Finance: An Introduction in Discrete Time.* (Third Edition.) Walter de Gruyter, Berlin.

Föllmer, H. and Weber, S. (2015). The axiomatic approach to risk measures for capital determination. *Annual Review of Financial Economics*, **7**, 301–337.

Furman, E. and Landsman, Z. (2006a). Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin: Journal of the International Actuarial Association*, **36**, 433–462.

Furman, E. and Landsman, Z. (2006b). On some risk-adjusted tail-based risk measures *Journal of Actuarial Practice*, **13**, 175–191.

Furman, E. and Zitikis, R. (2008). Weighted risk capital allocations. *Insurance: Mathematics and Economics* **42**, 459–465.

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

Giorgi, G.M. (1990). Bibliographic portrait of the Gini concentration ratio. *Metron*, **48**, 183–221.

Giorgi, G.M. (1993). A fresh look at the topical interest of the Gini concentration ratio. *Metron*, **51**, 83–98.

Grechuk, B., Molyboha, A. and Zabarankin, M. (2009). Maximum entropy principle with general deviation measures. *Mathematics of Operations Research*, **34**, 445–467.

Knight, J. and Satchel, S. (2001). *Return Distributions in Finance*. Butterworth–Heinemann, Oxford.

Landsman, Z.M. and Valdez, E.A. (2003). Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, **7**, 55–71.

Liu, F. and Wang, R. (2016). A theory for measures of tail risk. *Preprint available at SSRN:* `https://ssrn.com/abstract=2841909`.

Markowitz, H. (1991). *Portfolio Selection: Efficient Diversification of Investments* (Second edition). Wiley, New York.

McDonald, J.B. (1996). Probability distributions for financial models. In G.S. Maddala and C.R. Rao (eds.) *Handbook of Statistics*, **14**, *Statistical Methods in Finance*, Amsterdam, Elsevier Science, pp. 427–461.

Mao, T. and Wang, R. (2016). Risk aversion in regulatory capital principles. *Preprint available at SSRN:* `https://ssrn.com/abstract=2658669`.

McNeil, A.J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. (Revised Edition.) Princeton University Press, Princeton, NJ.

Kaina, M. and Rüschendorf, L. (2009). On convex risk measures on L^p-spaces. *Mathematical Methods in Operations Research*, **69**, 475–495.

Krätschmer, V., Schied, A. and Zähle, H. (2014). Comparative and quantitiative robustness for law-invariant risk measures. *Finance and Stochastics*, **18**, 271–295.

Panjer, H.H. (2001). Measurement of risk solvency requirements and allocation of capital within financial conglomerates. *Research Report 01-14.* Institute of Insurance and Pension Research, University of Waterloo, Canada.

Rachev, S. (Ed.) (2003). *Handbook of Heavy-tailed Distributions in Finance.* Elsevier Science, Amsterdam.

Rockafellar, R.T., Uryasev, S. and Zabarankin, M. (2006). Generalized deviation in risk analysis. *Finance and Stochastics*, **10**, 51–74.

Rockafellar, R. T., Uryasev, S. and Zabarankin, M. (2008). Risk tuning with generalized linear regression. *Mathematics of Operations Research*, **33**(3), 712–729.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, **2**(3), 225–243.

Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios.* Springer, Heidelberg.

Sandström, A. (2010). *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice.* Chapman and Hall, Boca Raton.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society*, **97**, 255–261.

Seal, H.L. (1980). Survival probabilities based on Pareto claim distributions. *ASTIN Bulletin: Journal of the International Actuarial Association*, **11**, 61–71.

Tsanakas, A. and Millossovich, P. (2016). Sensitivity analysis using risk measures. *Risk Analysis*, **36**(1), 30–48.

Yaari, M.E. (1987). The dual theory of choice under risk. *Econometrica*, **55**, 95–115.

Yitzhaki, S. and Schechtman, E. (2013). *The Gini Methodology: A Primer on a Statistical Methodology.* Springer, New York, NY.