

A FORM OF MULTIVARIATE PARETO DISTRIBUTION WITH APPLICATIONS TO FINANCIAL RISK MEASUREMENT

BY

JIANXI SU AND EDWARD FURMAN

ABSTRACT

A new multivariate distribution possessing arbitrarily parametrized and positively dependent univariate Pareto margins is introduced. Unlike the probability law of Asimit *et al.* (2010), the structure in this paper is absolutely continuous with respect to the corresponding Lebesgue measure. The distribution is of importance to actuaries through its connections to the popular frailty models, as well as because of the capacity to describe dependent heavy-tailed risks. The genesis of the new distribution is linked to a number of existing probability models, and useful characteristic results are proved. Expressions for, e.g., the decumulative distribution and probability density functions, (joint) moments and regressions are developed. The distributions of minima and maxima, as well as, some weighted risk measures are employed to exemplify possible applications of the distribution in insurance.

KEYWORDS

Multivariate Pareto distributions, characterizations, dependence, weighted risk measures, minima, maxima.

1. INTRODUCTION

At the outset, we fix the probability space $(\Omega, \Sigma, \mathbf{P})$ and define the random vector (r.v.) $\mathcal{X} \ni \mathbf{X} = (X_1, \dots, X_n)'$ as a map from (Ω, Σ) to the n ($\in \mathbf{N}$)-dimensional Borel space $(\mathbf{R}_+^n, \mathcal{B}(\mathbf{R}_+^n))$. The cumulative distribution function (c.d.f.) of \mathbf{X} is in the sequel denoted by $F_{1,\dots,n}(x_1, \dots, x_n) := \mathbf{P}[X_1 \leq x_1, \dots, X_n \leq x_n]$, and the corresponding probability density function (p.d.f.) by $f_{1,\dots,n}(x_1, \dots, x_n) := \partial^n / (\partial x_1 \cdots \partial x_n) F_{1,\dots,n}(x_1, \dots, x_n)$, where $(x_1, \dots, x_n)' \in \mathbf{R}_+^n := (0, \infty)^n$. Finally, $F_i(x)$ and $f_i(x)$ denote, respectively, the marginal c.d.f. and p.d.f. of X_i , $i = 1, \dots, n$. Clearly, when the coordinates of \mathbf{X} are stochastically independent, there is only one way to formulate the c.d.f. $F_{1,\dots,n}$, whereas the shapes of the just-mentioned c.d.f. are infinite otherwise.

We further discuss the so-called multivariate reduction approach to creating random vectors with dependent coordinates. This paves the way to introducing the main object of our interest in Section 2. For applications of the multivariate reduction method in insurance, we refer to, e.g., Vernic (1997, 2000), Pfeifer and Nešlehová (2004), Furman and Landsman (2005, 2010), Boucher *et al.* (2008) and Tsanakas (2008), as well as to the references therein.

Let $\mathcal{Y} \ni \mathbf{Y} = (Y_1, \dots, Y_{n+1})'$ be an $(n + 1)$ -variate r.v. with mutually independent univariate margins distributed gamma. Namely, for $j = 1, \dots, n + 1$, the p.d.f. of $Y_j \sim Ga(\gamma_j (\in \mathbf{R}_+), \alpha_j (\in \mathbf{R}_+))$ is given by

$$g(y; \gamma_j, \alpha_j) = e^{-\alpha_j y} \frac{y^{\gamma_j - 1} \alpha_j^{\gamma_j}}{\Gamma(\gamma_j)}, \quad y \in \mathbf{R}_+, \quad (1.1)$$

with the corresponding Laplace transform being well-defined on $\mathbf{R}_+ := (0, \infty)$ (the interval of interest herein) and given by

$$\hat{G}(x; \gamma_j, \alpha_j) = \int_0^\infty e^{-xy} g(y) dy = \left(1 + \frac{x}{\alpha_j}\right)^{-\gamma_j}. \quad (1.2)$$

Definition 1.1 (Furman, 2008; Furman and Landsman, 2010). *Let $A \in Mat_{n \times (n+1)}(\mathbf{R}_{0,+})$ denote a deterministic $n \times (n + 1)$ matrix with suitable non-negative entries. Then $\mathbf{X} = A\mathbf{Y}$ is distributed n -variate gamma with shape parameters $\gamma_i^* = t_i(\gamma_1, \dots, \gamma_{n+1})$ and rate parameters $\alpha_i^* = u_i(\alpha_1, \dots, \alpha_{n+1})$ for appropriate positive Borel functions $t_i(\cdot), u_i(\cdot), i = 1, \dots, n$. Succinctly, we write $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}^*, \boldsymbol{\alpha}^*)$, where $\boldsymbol{\gamma}^* = (\gamma_1^*, \dots, \gamma_n^*)' \in \mathbf{R}_+^n$ and $\boldsymbol{\alpha}^* = (\alpha_1^*, \dots, \alpha_n^*)' \in \mathbf{R}_+^n$ are vectors of parameters.*

Example 1.1 (Mathai and Moschopoulos, 1991; see also Cherian, 1941 and Ramabhadran, 1951). *Let $Y_j \sim Ga(\gamma_j, \alpha_j), j = 1, \dots, n + 1$ be mutually independent random variables distributed gamma with arbitrary parameters, and choose the matrix A such that, for $i = 1, \dots, n$ and $\sigma_i > 0$, it holds that $(A)_{i,n+1} = \alpha_{n+1}/\alpha_i, (A)_{i,i} \equiv 1$ and zero otherwise. Then $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}^*, \boldsymbol{\alpha})$, where $\boldsymbol{\gamma}^* = (\gamma_{n+1} + \gamma_1, \dots, \gamma_{n+1} + \gamma_n)'$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)'$ are two n -variate vectors of parameters.*

Example 1.2 (Mathai and Moschopoulos, 1992; see also, Furman, 2008). *Let $Y_j \sim Ga(\gamma_j, \alpha_j), j = 1, \dots, n$ be mutually independent random variables distributed gamma with arbitrary parameters, and choose the matrix A such that, for $i = 1, \dots, n, j = 1, \dots, i$ and $\sigma_i > 0$, it holds that $(A)_{i,j} = \alpha_j/\sigma_i$ and zero otherwise. Then $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}^*, \boldsymbol{\sigma})$, where $\boldsymbol{\gamma}^* = (\gamma_1^*, \dots, \gamma_n^*)', \gamma_i^* = \sum_{j=1}^i \gamma_j$ and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)'$ are two n -variate vectors of parameters.*

In the present paper, we employ the following modification of Examples 1.1 and 1.2.

Example 1.3 (Furman, 2008). *Let $Y_j \sim Ga(\gamma_j, \alpha_j), j = 1, \dots, n + 1$ be again mutually independent random variables distributed gamma with arbitrary param-*

eters, and choose the matrix A such that, for $i = 1, \dots, n$, $j = 1, \dots, i$ and $\sigma_i > 0$, it holds that $(A)_{i,j} = \alpha_j/\sigma_i$, $(A)_{i,n+1} = \alpha_{n+1}/\sigma_i$ and zero otherwise. Then $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}^*, \boldsymbol{\sigma})$, where $\boldsymbol{\gamma}^* = (\gamma_1^*, \dots, \gamma_n^*)'$, $\gamma_i^* = \gamma_{n+1} + \sum_{j=1}^i \gamma_j$ and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)'$ are two n -variate vectors of parameters.

In the sequel, we embark on the idea in Example 1.3 to introduce an encompassing yet tractable multivariate distribution with univariate margins distributed Pareto. We note in passing that a real-valued r.v. is said to be distributed Pareto of the 2nd kind, succinctly $X \sim Pa(II)(\mu, \sigma, \alpha)$, where $\mu \in \mathbf{R}$ is a location parameter, $\sigma \in \mathbf{R}_+$ is a scale parameter and $\alpha \in \mathbf{R}_+$ is a tail index, if its c.d.f. is given by

$$F_X(x; \mu, \sigma, \alpha) = 1 - \left(1 + \frac{x - \mu}{\sigma}\right)^{-\alpha}, \quad x > \mu, \tag{1.3}$$

(see, e.g., Arnold, 1983; Pareto, 1897; Kotz *et al.*, 2000). Similarly to Asimit *et al.* (2010), we set $\mu = 0$, which conveniently makes the support of the distribution coincide with the positive half of the real line, i.e., $supp_F = \{x \in \mathbf{R} : f(x) \neq 0\} = \mathbf{R}_+$ and does not lead to any loss of generality. The resulting distribution (Lomax distribution), notationally $Pa(II)(\sigma, \alpha)$, enjoys a great variety of applications in all areas of applied mathematics in general and in actuarial science in particular, as it naturally arises in the extreme value theory as the limiting distribution of the excess-of-loss r.v. $X_d := X - d | X > d$, where $d(\in \mathbf{R}_+)$ denotes a threshold (see, e.g., Pickands, 1975; Balkema and de Haan, 1974).

The rest of the paper is organized as follows. In Section 2, a multivariate probability structure with dependent Pareto-distributed univariate margins is introduced and linked to a number of existing multivariate models. Then distributional properties of the new structure are derived and some characterization results are proved in Sections 2 and 3. In Section 4, the new multivariate Pareto is reintroduced as a variant of the minima-based multiple risk factor model, and some applications to notions of actuarial interest are considered. In Section 5, an application of the model is elucidated with the help of a numerical example borrowed from the context of default risk. Section 6 concludes the paper. All proofs are relegated to the appendix to facilitate the reading.

2. NEW MULTIVARIATE PARETO DISTRIBUTION

Let $\mathbf{Y} = (Y_1, \dots, Y_{n+1})'$ be an r.v. with mutually independent coordinates $Y_j \sim Ga(\gamma_j, 1)$, $\gamma_j \in \mathbf{R}_+$, and choose the matrix $A_c \in Mat_{n \times (n+1)}$ such that $(A_c)_{i,j} = c_{i,j}/\sigma_i$, where $c_{i,j} \in \{0, 1\}$ are deterministic constants, $\sigma_i \in \mathbf{R}_+$, $i = 1, \dots, n$ and $j = 1, \dots, n + 1$. The following definition unifies Examples 1.1–1.3 and serves as an auxiliary tool for constructing the multivariate Pareto distribution of interest.

Definition 2.1. Let $\mathbf{X} = (X_1, \dots, X_n)' = A_c \mathbf{Y}$, then it follows an n -variate gamma distribution, notationally $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}_c^*, \boldsymbol{\sigma})$, where $\boldsymbol{\gamma}_c^* = (\gamma_{c,1}^*, \dots, \gamma_{c,n}^*)' \in \mathbf{R}_+^n$ with $\gamma_{c,i}^* = \sum_{j=1}^{n+1} c_{i,j} \gamma_j$, $i = 1, \dots, n$ and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)' \in \mathbf{R}_+^n$ are two vectors of parameters.

We note in passing that Definition 2.1 (auxiliary for the present paper) establishes an encompassing multivariate probability law with gamma-distributed univariate margins and an additive background risk dependence structure (see, Gollier and Pratt, 1996; Tsanakas, 2008; Furman and Landsman, 2010 for applications of the additive background risk models in economics and actuarial science). More specifically, the following simple special cases of $Ga_{1,\dots,n}(\boldsymbol{\gamma}_c^*, \boldsymbol{\sigma})$ readily recover the models of, respectively, Mathai and Moschopoulos (1991, 1992) and Furman (2008):

- $c_{i,i} = c_{i,n+1} \equiv 1$ for $i = 1, \dots, n$ and zero otherwise — Example 1.1;
- $c_{i,j} \equiv 1$ for $1 \leq j \leq i \leq n$ and zero otherwise — Example 1.2;
- $c_{i,j} = c_{i,n+1} \equiv 1$ for $1 \leq j \leq i \leq n$ and zero otherwise — Example 1.3.

Some elementary but useful properties of $\mathbf{X} \sim Ga_{1,\dots,n}$ follow directly by definition or from the Laplace transform that is established next.

Proposition 2.1. Let $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}_c^*, \boldsymbol{\sigma})$ be the r.v. distributed multivariate gamma as in Definition 2.1, then the corresponding Laplace transform is given by

$$\hat{G}_{1,\dots,n}(\mathbf{t}) = \prod_{j=1}^{n+1} \left(1 + \sum_{i=1}^n \frac{c_{i,j}}{\sigma_i} t_i \right)^{-\gamma_j}$$

and it is well defined on \mathbf{R}_+^n .

Immediate consequences of Proposition 2.1 are, for $k, l = 1, \dots, n$, that

- the distribution of $\mathbf{X} \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}_c^*, \boldsymbol{\sigma})$ is “marginally closed,” i.e., $X_k \sim Ga(\gamma_{c,k}^* (\in \mathbf{R}_+), \sigma_k (\in \mathbf{R}_+))$;
- the expectation of the k th coordinate is $\mathbf{E}[X_k] = \gamma_{c,k}^* / \sigma_k$;
- the variance of the k th coordinate is $\mathbf{Var}[X_k] = \gamma_{c,k}^* / \sigma_k^2$;
- for $k \neq l$, the covariance between the coordinates X_k and X_l is non-negative and given by

$$\mathbf{Cov}[X_k, X_l] = \frac{\sum_{j=1}^{n+1} c_{k,j} c_{l,j} \gamma_j}{\sigma_k \sigma_l}$$

- for $k \neq l$, the Pearson linear correlation between the coordinates X_k and X_l is non-negative and given by

$$\rho[X_k, X_l] = \frac{\sum_{j=1}^{n+1} c_{k,j} c_{l,j} \gamma_j}{\sqrt{\gamma_{c,k}^* \gamma_{c,l}^*}}$$

We are now in a position to introduce the multivariate Pareto distribution of interest. In fact, simple observation (1.2) along with Proposition 2.1 results in the following definition.

Definition 2.2. We call the r.v. $\mathbf{X} = (X_1, \dots, X_n)'$ having the decumulative distribution function (d.d.f.)

$$\bar{F}_{1,\dots,n}(x_1, \dots, x_n) = \prod_{j=1}^{n+1} \left(1 + \sum_{i=1}^n \frac{c_{i,j}}{\sigma_i} x_i \right)^{-\gamma_j}, \quad \text{where } (x_1, \dots, x_n)' \in \mathbf{R}_+^n, \tag{2.1}$$

a multivariate Pareto of the 2nd kind. Succinctly, we write $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$, where $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)'$, $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_n)'$ are two deterministic vectors of positive parameters, and $\gamma_{n+1} \in \mathbf{R}_+$ and $c \in \text{Mat}_{n \times (n+1)}(\{0, 1\})$ are scalar-valued and matrix-valued parameters, respectively.

Generally, distributions with Paretian tails have been applied in a multitude of areas. Herein we refer to: Benson *et al.* (2007) for applications in modeling catastrophic risk; Koedijk *et al.* (1990), Longin (1996), Gabaix *et al.* (2003) for applications in general financial phenomena; Cebrián *et al.* (2003) for applications in insurance pricing; and Soprano *et al.* (2010), Chavez-Demoulin *et al.* (2015) for applications in risk management.

Specifically, the probability law in Definition 2.2 is a generalization of the classical multivariate Pareto distribution of Arnold (1983) with the d.d.f. $\bar{F}_{1,\dots,n}^{\text{Arnold}}$. Indeed, set $c_{i,j} = c_{\bullet,j}$, $i = 1, \dots, n$, $j = 1, \dots, n + 1$ in (2.1) and obtain, for $\gamma_c^* = \sum_{j=1}^{n+1} c_{\bullet,j} \gamma_j$, that

$$(2.1) = \left(1 + \sum_{i=1}^n \frac{x_i}{\sigma_i} \right)^{-\gamma_c^*} = \bar{F}_{1,\dots,n}^{\text{Arnold}}(x_1, \dots, x_n), \quad \text{where } (x_1, \dots, x_n)' \in \mathbf{R}_+^n. \tag{2.2}$$

That being said, unlike the classical multivariate Pareto distribution of Arnold (1983), the structure in Definition 2.2 incorporates stochastic independence — set $c_{i,i} \equiv 1$, $i = 1, \dots, n$ and zero otherwise and obtain, for $\bar{F}_{1,\dots,n}^\Pi$ denoting the d.d.f. of a multivariate Pareto with independent margins, that

$$(2.1) = \prod_{i=1}^n \left(1 + \frac{x_i}{\sigma_i} \right)^{-\gamma_i} = \bar{F}_{1,\dots,n}^\Pi(x_1, \dots, x_n), \quad \text{where } (x_1, \dots, x_n)' \in \mathbf{R}_+^n.$$

Consequently, the new multivariate Pareto distribution meaningfully fills the gap between the multivariate probability distributions with independent and Arnold-dependent Pareto-distributed margins.

In addition, unlike (2.2), $Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ allows for distinct marginal tail indices (see, Proposition 2.2 below). Furthermore, the new multivariate Pareto distribution unifies the probability models studied recently in Chiragiev and

Landsman (2009). Namely, in order to obtain their “flexible Pareto type I and II” we choose $c_{i,i} = c_{i,n+1} \equiv 1, i = 1, \dots, n$ and zero otherwise and $c_{i,j} \equiv 1$ for $1 \leq j \leq i \leq n$ and zero otherwise, respectively.

Lastly, but perhaps most importantly in actuarial applications, d.d.f. (2.1) admits stochastic representations that mimic the multiplicative background risk model (Franke *et al.*, 2006) and the minima-based common shock model (Bowers *et al.*, 1997) (see, respectively, Theorems 2.2 and 4.1 in this paper). Stochastic representations are a very welcome facet, since they endow probabilistic models with an important feature of interpretability, and as a result contribute greatly to the process of model selection and implementation.

We further document several simple properties of the multivariate Pareto with d.d.f. (2.1). The proofs are straightforward and thus omitted.

Proposition 2.2. *Let $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then, for $i = 1, \dots, n$, the marginal d.d.f. of X_i is*

$$\bar{F}_i(x_i) = \left(1 + \frac{x_i}{\sigma_i}\right)^{-\gamma_{c,i}^*}, \quad x_i \in \mathbf{R}_+,$$

that is $X_i \sim Pa(II)(\sigma_i, \gamma_{c,i}^*)$, where $\gamma_{c,i}^* = \sum_{j=1}^{n+1} c_{i,j}\gamma_j$. Also, for $i = 1, \dots, n$ and setting $\gamma_{c,i}^* > 1$, we have that

$$E[X_i] = \sigma_i / (\gamma_{c,i}^* - 1), \tag{2.3}$$

and furthermore setting $\gamma_{c,i}^* > 2$, we obtain that

$$\text{Var}[X_i] = \sigma_i^2 \gamma_{c,i}^* / \left((\gamma_{c,i}^* - 1)^2 (\gamma_{c,i}^* - 2) \right). \tag{2.4}$$

In Proposition 2.2, the substitution $c_{i,i} = c_{i,n+1} \equiv 1, i = 1, \dots, n$ and zero otherwise yields Theorem 1 of Chiragiev and Landsman (2009), whereas the substitution $c_{i,j} \equiv 1$ for $1 \leq j \leq i \leq n$ and zero otherwise results in their Theorem 5.

In what follows, we develop an expression for the joint p.d.f. of the multivariate Pareto distribution of interest. To this end, let

$$\prod_{i=1}^n \sum_{j=1}^{n+1} c_{i,j} y_j = \sum_{i_j \in I} d_c(i_1, \dots, i_{n+1}) \prod_{j=1}^{n+1} y_j^{i_j}, \tag{2.5}$$

where I establishes a set of positive integer indices such that $\sum_{j=1}^{n+1} i_j = n$, and $d_c(i_1, \dots, i_n)$ are appropriately chosen constants. Also, let

$$(\gamma)_n = \frac{\Gamma(\gamma + n)}{\Gamma(\gamma)}, \quad \text{where } \gamma \in \mathbf{R}_+ \text{ and } n \in \mathbf{N}$$

denote the Pochhammer symbol.

Theorem 2.1. Let $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then the corresponding joint p.d.f. is formulated, for $(x_1, \dots, x_n)' \in \mathbf{R}_+^n$, as

$$f_{1,\dots,n}(x_1, \dots, x_n) = \sum_{\forall i_j \in I} d_c(i_1, \dots, i_{n+1}) \prod_{j=1}^{n+1} \frac{(\gamma_j)_{i_j}}{\prod_{l=1}^n \sigma_l} \left(1 + \sum_{i=1}^n c_{i,j} \frac{x_i}{\sigma_i} \right)^{-(\gamma_j+i_j)}, \tag{2.6}$$

where $d_c(i_1, \dots, i_{n+1})$ are appropriately chosen constants and $i_j \in I$.

In general, the constants $d_c(i_1, \dots, i_n)$ can be rather involved. For an insight, we show how (2.6) reduces to the p.d.f. of the classical multivariate Pareto distribution of Arnold (1983). To this end, set $c_{i,j} \equiv 1$ for $i = 1, \dots, n$ and $j = 1 \dots, n + 1$. Then from (2.5), we have that

$$d_c(i_1, \dots, i_{n+1}) = \binom{n}{i_1, \dots, i_{n+1}},$$

with the right-hand side denoting the multinomial coefficient. On the other hand, as (2.6) must integrate to one and since for Arnold’s multivariate Pareto distribution, we have, for $\gamma^* = \gamma_1 + \dots + \gamma_{n+1}$, that

$$\prod_{j=1}^{n+1} \left(1 + \sum_{i=1}^n c_{i,j} \frac{x_i}{\sigma_i} \right)^{-(\gamma_j+i_j)} = \left(1 + \sum_{i=1}^n \frac{x_i}{\sigma_i} \right)^{-(\gamma^*+n)},$$

we obtain

$$(2.6) = \frac{(\gamma^*)_n}{\prod_{i=1}^n \sigma_i} \left(1 + \sum_{i=1}^n \frac{x_i}{\sigma_i} \right)^{-(\gamma^*+n)}, \quad \text{for } (x_1, \dots, x_n)' \in \mathbf{R}_+^n,$$

as required.

The following theorem establishes a useful characteristic relation in the context of the multivariate Pareto distribution of interest, and it also plays an important role when deriving the formula for the corresponding Pearson linear correlation (see, Theorem 3.1 in Section 3). In the sequel “ $\stackrel{d}{=}$ ” denotes equality in distribution.

Theorem 2.2. Let $\boldsymbol{\Lambda} = (\Lambda_1, \dots, \Lambda_n)'$ be an r.v. with independent and exponentially distributed univariate margins $\Lambda_i \sim \text{Exp}(1)$, $i = 1, \dots, n$, and denote by $\boldsymbol{\Xi} = (\Xi_1, \dots, \Xi_n)' \sim Ga_{1,\dots,n}(\boldsymbol{\gamma}_c^*, \boldsymbol{\sigma})$ the n -variate gamma distribution introduced in Definition 2.1; here $\boldsymbol{\gamma}_c^* = (\gamma_{c,1}^*, \dots, \gamma_{c,n}^*)' \in \mathbf{R}_+^n$ with $\gamma_{c,i}^* = \sum_{j=1}^{n+1} c_{i,j} \gamma_j$, and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_n)' \in \mathbf{R}_+^n$ are vectors of parameters. Assume that $\boldsymbol{\Lambda}$ and $\boldsymbol{\Xi}$ are stochastically independent, then $\mathbf{X} = (X_1, \dots, X_n)'$ has d.d.f. (2.1), and it is thus the n -variate Pareto distribution introduced in Definition 2.2 if and only if $(X_1, \dots, X_n)' \stackrel{d}{=} (\Lambda_1/\Xi_1, \dots, \Lambda_n/\Xi_n)'$.

Theorem 2.2 establishes the multiplicative background risk representation of the multivariate probabilistic structure of main interest herein (see, Franke *et al.*, 2006; Meyers, 2007, Asimit *et al.*, 2013, 2016 for applications of the multiplicative background risk models in economics and actuarial science).

We conclude this section with yet another characterization of the multivariate Pareto distribution of interest and its two implications. Namely, let $\wedge_{i=1}^n X_i =: X_- \sim F_-$ and $\vee_{i=1}^n X_i =: X_+ \sim F_+$ denote, respectively, the minima and the maxima r.v.'s, and let $X_i \sim F_i, i = 1, \dots, n$ be univariate coordinates of the multivariate Pareto r.v. of interest in this paper.

Theorem 2.3. *Let $\mathbf{X} = (X_1, \dots, X_n)'$ be distributed $Pa_{1, \dots, n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as per Definition 2.2, then X_- admits the mixture representation as $X_- | \Lambda = \lambda \sim Exp(\lambda)$ and $\Lambda \stackrel{d}{=} Z_1 + \dots + Z_{n+1}$, where $Z_j, j = 1, \dots, n + 1$ are univariate mutually independent r.v.'s distributed gamma.*

An important corollary of Theorem 2.3 is a random parameter representation (see, e.g., Feller, 1966) of the d.d.f.'s of X_- and X_+ . The following lemma is crucial in studying the distribution of X_- in Theorem 2.4.

Lemma 2.1 (Moschopoulos, 1985; Furman and Landsman, 2005). *For $i = 1, \dots, n$, let $Z_i \sim Ga(\gamma_i(\in \mathbf{R}_+), \alpha_i(\in \mathbf{R}_+))$ denote independent gamma-distributed r.v.'s. Then the distribution of $Z = Z_1 + \dots + Z_n$ is gamma with a random shape parameter. More specifically, $Z \sim Ga(\gamma^* + K, \alpha_+)$, where $\gamma^* = \gamma_1 + \dots + \gamma_n, \alpha_+ = \vee_{i=1}^n \alpha_i$ and K is an integer-valued non-negative r.v. with the probability mass function (p.m.f.) given by*

$$p_k = \mathbf{P}[K = k] = c_+ \delta_k, \quad k = 0, 1, \dots, \tag{2.7}$$

where

$$c_+ = \prod_{i=1}^n \left(\frac{\alpha_i}{\alpha_+} \right)^{\gamma_i}, \quad \delta_0 = 1$$

and

$$\delta_k = k^{-1} \sum_{l=1}^k \sum_{i=1}^n \gamma_i \left(1 - \frac{\alpha_i}{\alpha_+} \right)^l \delta_{k-l} \quad \text{for } k > 0.$$

Theorem 2.4. *Let $\mathbf{X} \sim Pa_{1, \dots, n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then $X_- \sim Pa(II)(\alpha_+(\boldsymbol{\sigma}), \gamma^* + K)$, where $\alpha_+(\boldsymbol{\sigma}) = \vee_{j=1}^{n+1} \left(\sum_{i=1}^n \frac{c_{i,j}}{\sigma_i} \right)^{-1}$, K is an integer-valued r.v. with p.m.f. (2.7) and $\gamma^* = \gamma_1 + \dots + \gamma_{n+1} > 1$.*

Although Theorem 2.4 demonstrates that the minima r.v. X_- is distributed mixed Pareto with random tail index parameter, the next theorem asserts that the maxima r.v. X_+ has a d.d.f. that is a linear combination of the d.d.f.'s of such mixed Pareto-distributed r.v.'s. The proof is similar to the one of Proposition 2 in Vernic (2011) and is thus omitted.

Theorem 2.5. Assume that $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then the d.d.f. of the maxima r.v. is given by

$$\bar{F}_+(x) = \sum_{S \subseteq \{1,\dots,n\}} (-1)^{|S|-1} \bar{F}_{S^-}(x), \quad x \in \mathbf{R}_+, \tag{2.8}$$

where $X_{S^-} = \bigwedge_{s \in S} X_s$ and $X_{S^-} \sim F_{S^-}$.

3. BIVARIATE QUANTITIES OF INTEREST

It is worthwhile to make an additional observation before stating the main result of this section. Namely, we note in passing that for $1 \leq k \neq l \leq n$, an r.v. $(\Xi_k, \Xi_l)'$ distributed the bivariate gamma per Definition 2.1 and an $(n + 1)$ -variate r.v. $\mathbf{Y} = (Y_1, \dots, Y_{n+1})'$ having mutually independent coordinates $Y_j \sim Ga(\gamma_j, 1)$, $\gamma_j \in \mathbf{R}_+$, the following stochastic representation holds

$$(\sigma_k \Xi_k, \sigma_l \Xi_l)' \stackrel{d}{=} (Y_{c,(k,l)} + Y_{c,k}, Y_{c,(k,l)} + Y_{c,l})', \tag{3.1}$$

where $Y_{c,(k,l)} = \sum_{j=1}^{n+1} c_{k,j} c_{l,j} Y_j$, $Y_{c,k} = \sum_{j=1}^{n+1} c_{k,j} (1 - c_{l,j}) Y_j$ and $Y_{c,l} = \sum_{j=1}^{n+1} c_{l,j} (1 - c_{k,j}) Y_j$ are mutually independent gamma-distributed r.v.'s with the shape parameters $\gamma_{c,(k,l)} = \sum_{j=1}^{n+1} c_{k,j} c_{l,j} \gamma_j$, $\gamma_{c,k} = \sum_{j=1}^{n+1} c_{k,j} (1 - c_{l,j}) \gamma_j$ and $\gamma_{c,l} = \sum_{j=1}^{n+1} c_{l,j} (1 - c_{k,j}) \gamma_j$, respectively.

We next show that the covariance of a random couple within the multivariate Pareto of interest in this paper can be formulated using the $(q + 1) \times q$ hypergeometric function (see, Gradshteyn and Ryzhik, 2007), which is formulated as

$${}_{q+1}F_q(a_1, \dots, a_{q+1}; b_1, \dots, b_q; z) := \sum_{k=0}^{\infty} \frac{(a_1)_k, \dots, (a_{q+1})_k}{(b_1)_k, \dots, (b_q)_k} \frac{z^k}{k!}, \tag{3.2}$$

where $q \in \mathbf{Z}_+$. For a_1, \dots, a_{q+1} all positive, and these are the cases of interest in the present paper, the radius of convergence of the series is the open disk $|z| < 1$. On the boundary $|z| = 1$, the series converges absolutely if $h := b_1 + \dots + b_q - a_1 - \dots - a_{q+1} > 0$, and it converges except at $z = 1$ if $0 \geq h > -1$.

Theorem 3.1. Let $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2 and assume that both $\gamma_{c,k}^*$ and $\gamma_{c,l}^*$ exceed 2, then, for $0 \leq k \neq l \leq n$,

$$\text{Cov}[X_k, X_l] = \sigma_k \sigma_l \frac{1}{(\gamma_{c,k}^* - 1)(\gamma_{c,l}^* - 1)} ({}_3F_2(\gamma_{c,(k,l)}, 1, 1; \gamma_{c,k}^*, \gamma_{c,l}^*; 1) - 1).$$

An immediate consequence of Theorem 3.1 is that the maximal attainable Pearson correlation in the context of the multivariate Pareto distribution introduced in the present paper is not one. This consequence is however solely a result of the fact that the Pearson index of correlation exists only if the involved second

moments are finite, a pitfall that is well-known to non-life actuaries, which often deal with heavy-tailed losses (see, Embrechts *et al.*, 2002).

Corollary 3.1. *Let $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ and assume that both $\gamma_{c,k}^*$ and $\gamma_{c,l}^*$ exceed 2 for $0 \leq k \neq l \leq n$, then, for the Pearson correlation, it holds that $\text{Corr}[X_k, X_l] \in [0, 1/2)$.*

Another consequence of Theorem 3.1 pertains to two special cases of the multivariate Pareto introduced in this paper, and it is formulated as the following corollary.

Corollary 3.2. *Let $\mathbf{X}_1 = (X_{1,1}, \dots, X_{1,n})' \sim Pa_{1,\dots,n}^{(I)}$ and $\mathbf{X}_2 = (X_{2,1}, \dots, X_{2,n})' \sim Pa_{1,\dots,n}^{(II)}$ be distributed, respectively, the multivariate flexible Pareto of type I and II of Chiragiev and Landsman (2009). Then the corresponding covariances are readily obtained, for $0 \leq k \neq l \leq n$, as*

$$\text{Cov}[X_{1,k}, X_{1,l}] = \frac{\sigma_k \sigma_l}{(\gamma_k + \gamma_{n+1} - 1)(\gamma_l + \gamma_{n+1} - 1)} \times ({}_3F_2(\gamma_{n+1}, 1, 1; \gamma_k + \gamma_{n+1}, \gamma_l + \gamma_{n+1}; 1) - 1), \quad (3.3)$$

for $\gamma_k + \gamma_{n+1} > 2$ and $\gamma_l + \gamma_{n+1} > 2$, and as

$$\text{Cov}[X_{2,k}, X_{2,l}] = \frac{\sigma_k \sigma_l}{(\gamma_{c,k}^* - 1)(\gamma_{c,l}^* - 1)(\gamma_{c,l}^* - 2)}, \quad (3.4)$$

for $\gamma_{c,k}^* = \sum_{j=1}^k \gamma_j > 2$ and $\gamma_{c,l}^* = \sum_{j=1}^l \gamma_j > 2$.

We note in passing that expression (3.4) confirms the one derived in Chiragiev and Landsman (2009), whereas formula (3.3) complements the discussion therein. Also, the covariance of two r.v.'s coming from Arnold's multivariate Pareto distribution (see, Arnold, 1983) is readily obtained from both (3.3) and (3.4). More specifically, we set $\gamma_k = \gamma_l \equiv 0$ for all $0 \leq k \neq l \leq n$ and verify that (3.3) reduces to

$$\frac{\sigma_k \sigma_l}{(\gamma_{n+1} - 1)(\gamma_{n+1} - 1)} ({}_2F_1(1, 1; \gamma_{n+1}; 1) - 1) = \frac{\sigma_k \sigma_l}{(\gamma_{n+1} - 1)^2(\gamma_{n+1} - 2)},$$

for $\gamma_{n+1} > 2$. The verification is straightforward in the case of (3.4).

Generalized hypergeometric function (3.2) plays an important role when deriving the centered regression function $r(y) = \mathbf{E}[X - \mathbf{E}[X] | Y = y]$, where $y \in \mathbf{R}_+$ (see, Furman and Zitikis, 2008b, 2010, for applications of the function in insurance and finance). We next present the conditional d.d.f., followed by the centered regression function for a pair of r.v.'s having the probability law as in Definition 2.2. To this end, let

$$m(x) = \frac{\sigma_k}{\gamma_{c,(k,l)}} \left(1 + \frac{x}{\sigma_l} \right), \quad x \in \mathbf{R}_+.$$

Theorem 3.2. Let $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then the d.d.f. of X_k given $X_l = x_l, 0 \leq k \neq l \leq n$, is formulated as

$$\begin{aligned} \mathbf{P}[X_k > x_k | X_l = x_l] &= \left(\frac{\gamma_{c,(k,l)}}{\gamma_{c,l}^*} + \frac{\gamma_{c,l}}{\gamma_{c,l}^*} \left(1 + \frac{x_k}{\gamma_{c,(k,l)}m(x_l)} \right) \right) \left(1 + \frac{x_k}{\sigma_k} \right)^{-\gamma_{c,k}} \\ &\times \left(1 + \frac{x_k}{\gamma_{c,(k,l)}m(x_l)} \right)^{-\gamma_{c,(k,l)}-1}, \end{aligned} \tag{3.5}$$

where $x_k, x_l \in \mathbf{R}_+$.

Theorem 3.3. Let $\mathbf{X} \sim Pa_{1,\dots,n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then the centred regression function of X_k on $X_l, 0 \leq k \neq l \leq n$, is given, for $\gamma_{c,k}^* > 1$, by

$$r_k(x_l) = m(x_l) \sum_{i=1}^2 a_i {}_2F_1 \left(\gamma_{c,k}, 1; \gamma_{c,k}^* + 2 - i; -\frac{x_l}{\sigma_l} \right) - \sigma_k / (\gamma_{c,k}^* - 1), \tag{3.6}$$

where

$$a_1 = \frac{\gamma_{c,(k,l)}^2}{\gamma_{c,k}^* \gamma_{c,l}^*}, \quad a_2 = \frac{\gamma_{c,l} \gamma_{c,(k,l)}}{\gamma_{c,l}^* (\gamma_{c,k}^* - 1)} \quad \text{and } x_l \in \mathbf{R}_+.$$

The centred regression function is monotonically-increasing and concave.

We reiterate that our results readily recover the ones derived in Landsman and Chiragiev (2009). More specifically, by a simple alignment of notation in Theorem 3.3 above, we obtain Theorem 3 in loc. cit., whereas by choosing $\gamma_{c,k} = 0$ in Theorem 3.3 and hence for

$${}_2F_1 \left(0, 1; \gamma_{c,k}^* + 2 - i; -\frac{x_l}{\sigma_l} \right) = 1, \quad a_1 = \frac{\gamma_{c,k}^*}{\gamma_{c,l}^*} \quad \text{and } a_2 = \frac{(\gamma_{c,l}^* - \gamma_{c,k}^*)\gamma_{c,k}^*}{\gamma_{c,l}^* (\gamma_{c,k}^* - 1)},$$

we end up with Theorem 7 therein.

Clearly, the centred regression function of the new multivariate Pareto distribution is not linear, whereas it is well-known that the classical multivariate Pareto has linear regression (see, Arnold, 1983). Theorem 3.3 confirms the latter fact by setting $\gamma_{c,(k,l)} = \gamma_{c,k}^* = \gamma_{c,l}^* \equiv \gamma_{n+1}$ and $\gamma_{c,l} = \gamma_{c,k} \equiv 0$ in (3.6), which then reduces to the following linear form:

$$r_k(x_l) = \frac{\sigma_k}{\sigma_l} \left(x_l - \frac{\sigma_l}{\gamma_{n+1}(\gamma_{n+1} - 1)} \right),$$

for $\gamma_{n+1} > 1, 0 \leq k \neq l \leq n$ and $x_l \in \mathbf{R}_+$.

4. APPLICATIONS TO INSURANCE

In what follows, we assume that $\mathbf{X} = (X_1, \dots, X_n)'$ denotes a risk portfolio (r.p.) with $X_i, i = 1, \dots, n$ representing its risk components (r.c.'s). According to Theorem 2.2, if $\mathbf{X} \sim Pa_{1, \dots, n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$, then it admits the multiplicative background risk representation (see, Franke *et al.*, 2006; Meyers, 2007; Asimit *et al.*, 2013, 2016).

We next show that the new multivariate Pareto distribution can also be interpreted as a variant of the classical minima-based common shock model (see, e.g., Bowers *et al.*, 1997). To this end, assume that the i th r.c. of the r.p. is exposed to the set $\mathcal{R}_i = \{r \in \mathbf{N} : r \leq (n + 1)\}, i = 1, \dots, n$ of risk factors (r.f.'s) and let the r.v. $\mathbf{Y} = (Y_1, \dots, Y_{n+1})'$ stipulate the randomness of actuarial interest associated with the r.f.'s. The following theorem establishes the minima-based multiple risk factor representation of the multivariate Pareto proposed in the present paper. We note in passing that “ $*$ ” stands for the mixture operator, i.e., given two appropriately jointly measurable r.v.'s $X_\lambda \sim C(\cdot; \lambda)$ and $\Lambda \sim H$, it holds that $X_\lambda * \Lambda \stackrel{d}{=} X_\Lambda$.

Theorem 4.1. *Let $\mathbf{W}_i = (W_{i,1}, \dots, W_{i,n+1})'$ be r.v.'s with independent exponentially distributed margins $W_{i,j} \sim Exp(\lambda_{i,j}(\in \mathbf{R}_+))$, $i = 1, \dots, n$, $j = 1, \dots, n + 1$, and let A_c be a deterministic matrix of zero-one coefficients. Also, let $\Lambda = (\Lambda_1, \dots, \Lambda_{n+1})'$ be an r.v. having independent gamma-distributed margins with arbitrary shape parameters $\gamma_j(\in \mathbf{R}_+)$ and rate parameters equal to 1, $j = 1, \dots, n + 1$. Set, for $\sigma_i \in \mathbf{R}_+$ and $i = 1, \dots, n$,*

$$X_i = \sigma_i \bigwedge_{j=1, c_{i,j} \neq 0}^{n+1} (W_{i,j} * \Lambda_j), \tag{4.1}$$

then $\mathbf{X} = (X_1, \dots, X_n)' \sim Pa_{1, \dots, n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$.

Theorem 4.1 suggests that the multivariate Pareto distribution proposed in the present paper might be an appropriate formal framework for modeling-dependent default, survival or failure times when these times are exponentially distributed with random parameters. We elaborate on this observation in Section 5.

4.1. Actuarial risk measurement

Regulatory accords around the globe require that insurance companies carry out a careful assessment of their future losses. From now on, the r.v. $X : \Omega \rightarrow \mathbf{R}_+$ is interpreted as an insurance loss r.v., and \mathcal{X} denotes the collection of such r.v.'s.

Definition 4.1. *A risk measure is a functional map $H : \mathcal{X} \rightarrow [0, \infty]$.*

The literature on risk measures is vast and growing quickly. The following two indices are arguably the most popular amongst practitioners.

Definition 4.2. Let $X \in \mathcal{X}$ and fix $q \in [0, 1)$, then the Value-at-Risk (VaR) and the Conditional Tail Expectation (CTE) risk measures are, respectively, given by

$$VaR_q[X] = \inf\{x \in \mathbf{R} : \mathbf{P}[X \leq x] \geq q\}, \tag{4.2}$$

and

$$CTE_q[X] = \mathbf{E}[X | X > VaR_q[X]]. \tag{4.3}$$

We note in passing that both VaR and CTE are distorted as well as weighted risk measures (see, respectively, Wang, 1996; and Furman and Zitikis, 2008a).

Definition 4.3. (Furman and Zitikis (2008a), see also Choo and de Jong (2009, 2010.)) Let $w : \mathbf{R} \rightarrow \mathbf{R}_+$ be a non-decreasing Borel (weight) function such that $0 < \mathbf{E}[w(X)] < \infty$, then the class of weighted risk measures is defined as

$$\pi_w[X] = \frac{\mathbf{E}[Xw(X)]}{\mathbf{E}[w(X)]} \text{ for } X \in \mathcal{X}. \tag{4.4}$$

Let $v_1, w_1 : \mathbf{R}^n \rightarrow \mathbf{R}_+$ be two legitimate weight functions such that all expectations in (4.5) are finite and consider a generalized variant of (4.4)

$$\pi_{v_1, w_1}[\mathbf{\Lambda}] = \frac{\mathbf{E}[v_1(\mathbf{\Lambda})w_1(\mathbf{\Lambda})]}{\mathbf{E}[w_1(\mathbf{\Lambda})]} \text{ for } \mathbf{\Lambda} : \Omega \rightarrow \mathcal{R} \subseteq \mathbf{R}^n. \tag{4.5}$$

Proposition 4.1. Let $X|\mathbf{\Lambda} = \lambda \sim C(\cdot; \lambda(\in \mathcal{R} \subseteq \mathbf{R}^n))$ and assume that $\mathbf{\Lambda} \sim H_{1, \dots, n}$, then for any legitimate weight function, the functional $\pi_w[X]$ admits representation (4.5).

Corollary 4.1. Let $\mathbf{X} \sim Pa_{1, \dots, n}^c(\sigma, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then, for $i = 1, \dots, n$ and $q \in [0, 1)$, we have that

$$VaR_q[X_i] = \sigma_i \left((1 - q)^{-1/\gamma_{c,i}^*} - 1 \right).$$

Corollary 4.2. Let $X|\mathbf{\Lambda} = \lambda \sim C(\cdot; \lambda(\in \mathcal{R} \subseteq \mathbf{R}^n))$ and assume that $\mathbf{\Lambda} \sim H_{1, \dots, n}$, then the CTE risk measure of $X \sim F$ is, if exists and for $q \in [0, 1)$, given by

$$CTE_q[X] = \frac{\mathbf{E}[\bar{C}(VaR_q[X]; \mathbf{\Lambda})CTE_{q^*}[X|\mathbf{\Lambda}]]}{\bar{F}(VaR_q[X])} \text{ for } X \in \mathcal{X},$$

where $q^* = C(VaR_q[X]; \lambda)$.

Corollary 4.3. Let $\mathbf{X} \sim Pa_{1, \dots, n}^c(\sigma, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, then we have that the CTE risk measure is, if exists and for $i = 1, \dots, n$ and $q \in [0, 1)$,

given by

$$\begin{aligned}
 CTE_q[X_i] &= \mathbf{E}[X_i] \frac{\overline{F}_{X_i^*}(VaR_q[X_i])}{1 - q} + VaR_q[X_i], \\
 &= \mathbf{E}[X_i] + VaR_q[X_i] \frac{\gamma_{c,i}^*}{\gamma_{c,i}^* - 1},
 \end{aligned}
 \tag{4.6}$$

where $X_i^* \sim Pa(II)(\sigma_i, \gamma_{c,i}^* - 1)$.

The minima r.v. $X_- = \wedge_{i=1}^n X_i$ plays an important role in insurance mathematics (recall, e.g., the joint life policies in life insurance), as well as in general finance (think of, e.g., the first-to-default baskets).

Recall that the r.v. K has been defined as an integer-valued non-negative r.v. with the following p.m.f.

$$p_k = \mathbf{P}[K = k] = c_+ \delta_k, \quad k = 0, 1, \dots,
 \tag{4.7}$$

where

$$c_+ = \prod_{i=1}^n \left(\frac{\alpha_i}{\alpha_+} \right)^{\gamma_i}, \quad \delta_0 = 1,$$

and

$$\delta_k = k^{-1} \sum_{l=1}^k \sum_{i=1}^n \gamma_i \left(1 - \frac{\alpha_i}{\alpha_+} \right)^l \delta_{k-l} \quad \text{for } k > 0.$$

Proposition 4.2. *In the context of the multivariate Pareto of interest, the CTE risk measure of the minima can be written, if finite and for $q \in [0, 1)$, as*

$$CTE_q[X_-] = \mathbf{E}[X_-] \frac{\overline{F}_{X_-^*}(VaR_q[X_-])}{1 - q} + VaR_q[X_-],$$

where $X_-^* \sim Pa(\alpha_+(\boldsymbol{\sigma}), \gamma^* + Q - 1)$, where $\alpha_+(\boldsymbol{\sigma}) = \sqrt[n+1]{\sum_{i=1}^n \frac{c_{i,j}}{\sigma_i}}^{-1}$, $\gamma^* = \gamma_1 + \dots + \gamma_{n+1}$ and Q is an integer-valued r.v. with the p.m.f. obtained from the p.m.f. of K with the help of the following change of measure

$$q_k = \frac{1}{\mathbf{E}[X_-]} \frac{\alpha_+(\boldsymbol{\sigma})}{\gamma^* + k - 1} p_k, \quad k = 0, 1, \dots
 \tag{4.8}$$

Proposition 4.3. *In the context of the multivariate Pareto of interest, the CTE risk measure of the maxima can be written, if finite and for $q \in [0, 1)$, as the following linear combination:*

$$CTE_q[X_+] = \frac{1}{1 - q} \sum_{S \subseteq \{1, \dots, n\}} (-1)^{|S|-1} \mathbf{E}[X_{S-} | X_{S-} > VaR_q(X_+)] \overline{F}_{S-}(VaR_q(X_+)),$$

where $X_{S-} = \wedge_{s \in S \subseteq \{1, \dots, n\}} X_s$ and $X_{S-} \sim F_{S-}$.

Definition 4.4 (Furman and Zitikis, 2008b). Let $w : \mathbf{R} \rightarrow \mathbf{R}_+$ be a non-decreasing Borel function such that $0 < \mathbf{E}[w(Y)] < \infty$, then the functional $\Pi : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty]$ is referred to as the economic risk measure. Moreover, the special form of Π , given by

$$\Pi_w[X, Y] = \frac{\mathbf{E}[Xw(Y)]}{\mathbf{E}[w(Y)]} \quad \text{for } X \in \mathcal{X} \text{ and } Y \in \mathcal{X}, \tag{4.9}$$

is called a weighted economic risk measure.

We further derive an expression for the economic CTE risk measure, which is a particular case of (4.9) with $w(y) = \mathbf{1}\{y > VaR_q[Y]\}$, $q \in [0, 1)$ and $y \in \mathbf{R}_+$. To this end, we find the next proposition useful. The proof is plain and thus omitted.

Proposition 4.4. Let $\mathbf{X} \sim Pa_{1, \dots, n}^c(\boldsymbol{\sigma}, \boldsymbol{\gamma}, \gamma_{n+1})$ as in Definition 2.2, the d.d.f. of X_k given $X_l > x_l$, $0 \leq k \neq l \leq n$, is formulated as

$$\mathbf{P}[X_k > x_k | X_l > x_l] = \left(1 + \frac{x_k}{\sigma_k}\right)^{-\gamma_{c,k}} \left(1 + \frac{x_k}{\gamma_{c,(k,l)} m(x_l)}\right)^{-\gamma_{c,(l,k)}},$$

where

$$m(x_l) = \frac{\sigma_k}{\gamma_{c,(k,l)}} \left(1 + \frac{x_l}{\sigma_l}\right),$$

and x_k, x_l are both in \mathbf{R}_+ .

Proposition 4.5. In the context of the multivariate Pareto of interest, the economic CTE risk measure is given, for $\gamma_{c,k}^* > 1$, $q \in [0, 1)$ and $1 \leq k \neq l \leq n$, by

$$\mathbf{E}[X_k | X_l > VaR_q[X_l]] = \frac{\sigma_k}{\gamma_{c,k}^* - 1} {}_2F_1 \left(\gamma_{c,(k,l)}, 1; \gamma_{c,k}^*; \frac{VaR_q[X_l]}{\sigma_l + VaR_q[X_l]} \right).$$

To summarize, so far we have introduced and studied a new multivariate probability distribution with the univariate margins distributed Pareto of the second kind. The dependence structure of the new distribution is driven by a number of stochastic representations that are variants of the multiplicative background risk and the minima-based common shock models. We next employ the latter interpretation of the proposed multivariate probability distribution to exemplify its possible application to modeling and measuring default risk.

5. NUMERICAL ILLUSTRATION

For the sake of the discussion in this section, we adopt the view of the Financial Stability Board and the International Monetary Fund that the systemic risk can be caused by impairment of all or parts of the financial system, and more formally, we call the risk factor $j \in \{1, \dots, n + 1\}$ “systemic,” if $c_{i,j} = 1$

for at least two distinct r.c.'s $i \in \{1, \dots, n\}$. Similarly, we call the risk factor $j \in \{1, \dots, n + 1\}$ "idiosyncratic," if $c_{i,j} = 1$ for only one risk component $i \in \{1, \dots, n\}$.

Consider obligors in a default risk portfolio, each of which is exposed to exactly two distinct categories of fatal risk factors, e.g., systemic (category A) and idiosyncratic (category B). We assume that the risk factors from distinct risk categories are independent and that the hitting times (or occurrences) of defaults of the r.c.'s are exponentially distributed with random parameters distributed gamma. In fact, the future lifetime r.v. of the i th r.c. has exponential distribution with the random parameter $\sigma_i^{-1} \sum_{j=1}^{n+1} c_{i,j} \Lambda_j$, where Λ_j are distributed gamma with unit rate parameters, and i is 1, 2 or 3. Then Theorem 4.1 readily implies that the joint default times of the aforementioned r.c.'s has d.d.f. (2.1).

To illustrate the effect of the dependence structure on the joint default probability, we further set the dimension to $n = 3$ and specialize the setup above along the lines in Section 16.8 of Engelmann and Rauhmeier (2011) as well as employing the 2014's Annual Global Corporate Default Study and Rating Transitions of Standard & Poor's. (Standard & Poor's, 2015). More specifically, we set $\mu := E[\Lambda_j] \equiv 1.67$, fix the time horizon to 15 years and choose the corresponding default probability, p say, to be equal to 0.3198 (on par with the "B" credit rating of speculative entities). This yields the multivariate probability structure of Definition 2.2 with identically distributed margins having the parameters $\sigma_i \equiv \sigma = 122.39$ and $\gamma_{c,i}^* \equiv 3.33$, for $i = 1, 2$ and 3 .

Then we explore three different exposures of the obligors to the systemic and idiosyncratic r.f.'s. The distinct exposures are stipulated by appropriate choices of the c parameters gathered by matrices $A_c^{(k)}$, $k = 1, 2, 3$. We compare the aforementioned three exposures with the reference case in which no systemic risk is presented that is the joint d.d.f. of default times is a trivariate Pareto with independent margins. We note in passing that the expressions for the d.d.f.'s below readily follow from Theorem 4.1, whereas the values of the Pearson correlation coefficient are in non-trivial cases obtained with the help of Theorem 3.1.

Case (1). Only the systemic risk is presented, and all risk components are exposed to it. The exposure is represented schematically with the use of the following matrix, in which the rows and the columns represent r.c.'s and r.f.'s, respectively

$$A_c^{(1)} = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{array} \right).$$

The joint d.d.f. of the risk components is given by

$$\bar{F}^{(1)}(x_1, x_2, x_3) = \left(1 + \frac{x_1 + x_2 + x_3}{\sigma} \right)^{-2\mu},$$

where x_1, x_2, x_3 are all in \mathbf{R}_+ . This is obviously the d.d.f. of the classical trivariate Pareto distribution (Arnold, 1983). In this r.p., the Pearson correlation coefficient between any two of the r.c.'s is 0.3.

In the following two cases, both the systemic and idiosyncratic risks are presented.

Case (2). There are overall three uncorrelated idiosyncratic risk factors and one systemic risk factor. The exposure is gathered by the following block matrix:

$$A_c^{(2)} = \left(\begin{array}{c|cc} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array} \right).$$

The joint d.d.f. of the risk components is given by

$$\begin{aligned} \bar{F}^{(2)}(x_1, x_2, x_3) &= \left(1 + \frac{x_1 + x_2 + x_3}{\sigma} \right)^{-\mu} \left(1 + \frac{x_1}{\sigma} \right)^{-\mu} \left(1 + \frac{x_2}{\sigma} \right)^{-\mu} \left(1 + \frac{x_3}{\sigma} \right)^{-\mu}, \end{aligned}$$

where x_1, x_2, x_3 are all in \mathbf{R}_+ . This case corresponds to the “flexible Pareto type I” of Landsman and Chiragiev (2009). In this r.p., the Pearson correlation coefficient between any two of the r.c.'s is 0.09.

Case (3). The systemic risk is represented by two distinct risk factors of which one targets the entire risk portfolio and the other only hits r.c.'s #1 and #2. There is one idiosyncratic risk factor, and only r.c. #3 is exposed to it. The exposure block matrix is given by

$$A_c^{(3)} = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array} \right).$$

The joint d.d.f. of the risk components is

$$\begin{aligned} \bar{F}^{(3)}(x_1, x_2, x_3) &= \left(1 + \frac{x_1 + x_2 + x_3}{\sigma} \right)^{-\mu} \left(1 + \frac{x_1 + x_2}{\sigma} \right)^{-\mu} \left(1 + \frac{x_3}{\sigma} \right)^{-\mu}, \end{aligned}$$

where x_1, x_2, x_3 are all in \mathbf{R}_+ . In this r.p., the Pearson correlation coefficient between r.c. #1 and #2 is 0.3, and it is equal to 0.09 otherwise.

5.1. Expected times of the first default

The left panel of Figure 1 depicts the values of $CTE_q[X_-]$ for $q \in [0, 1)$, $X_- \in \mathcal{X}$ and portfolios (1) to (3) as well as the reference portfolio, denoted by (\perp). As

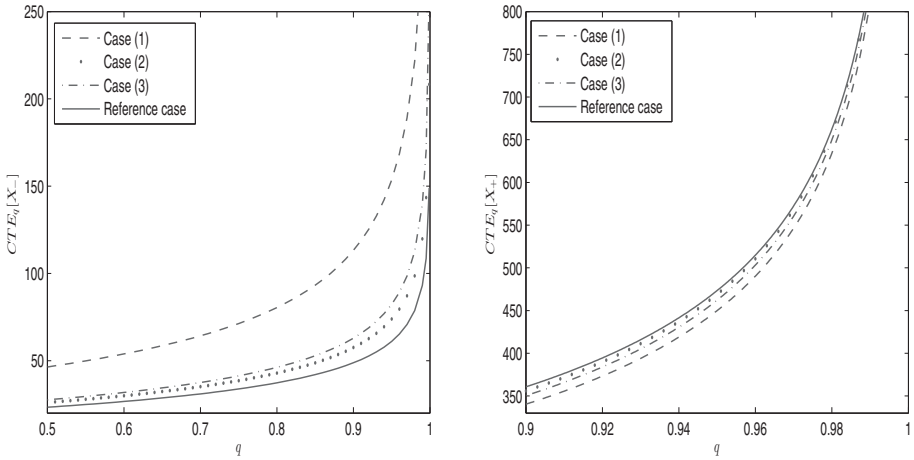


FIGURE 1: Conditional expected times of first (left panel) and last (right panel) default for portfolios (1)–(3) and the reference portfolio (\perp) for “B” rating r.p.’s with the probability of default $p = 0.3198$ and $\mu = 1.67$. Proposition 4.2 is employed to compute the values of CTE_q for $q \in [0, 1)$. (Color online)

the risk components are identically distributed, it is not difficult to see that the following ordering holds:

$$\bar{F}_-^{(1)} \geq_{st} \bar{F}_-^{(3)} \geq_{st} \bar{F}_-^{(2)} \geq_{st} \bar{F}_-^{(\perp)}, \tag{5.1}$$

where “ \geq_{st} ” denotes first-order stochastic dominance (FSD). Furthermore, since the CTE risk measure is known to preserve the FSD ordering, we also have that

$$CTE_q^{(1)}[X_-] \geq CTE_q^{(3)}[X_-] \geq CTE_q^{(2)}[X_-] \geq CTE_q^{(\perp)}[X_-],$$

for all $q \in [0, 1)$ and $X_- \in \mathcal{X}$. This conforms with Figure 1 (left panel), which hints that the r.p.’s with more significantly correlated r.c.’s enjoy higher, and thus more favorable, occurrence times of the first default.

The downside of high correlations is elucidated in Figure 2, in which we leave the probability of default p to be equal to 0.3198 (“B” rating) but vary the μ parameter that stipulates the effect of the risk factors. In this respect, we observe that the r.p.’s with stronger correlations between r.c.’s are more sensitive to the changes in the μ parameter, and therefore such r.p.’s must be monitored and stress tested more frequently.

5.2. Expected times of the last default

Figure 1 (right panel) depicts the values of $CTE_q[X_+]$ for $q \in [0, 1)$, $X_+ \in \mathcal{X}$ and portfolios (1) to (3) as well as the reference portfolio (\perp). Evoking Theorem

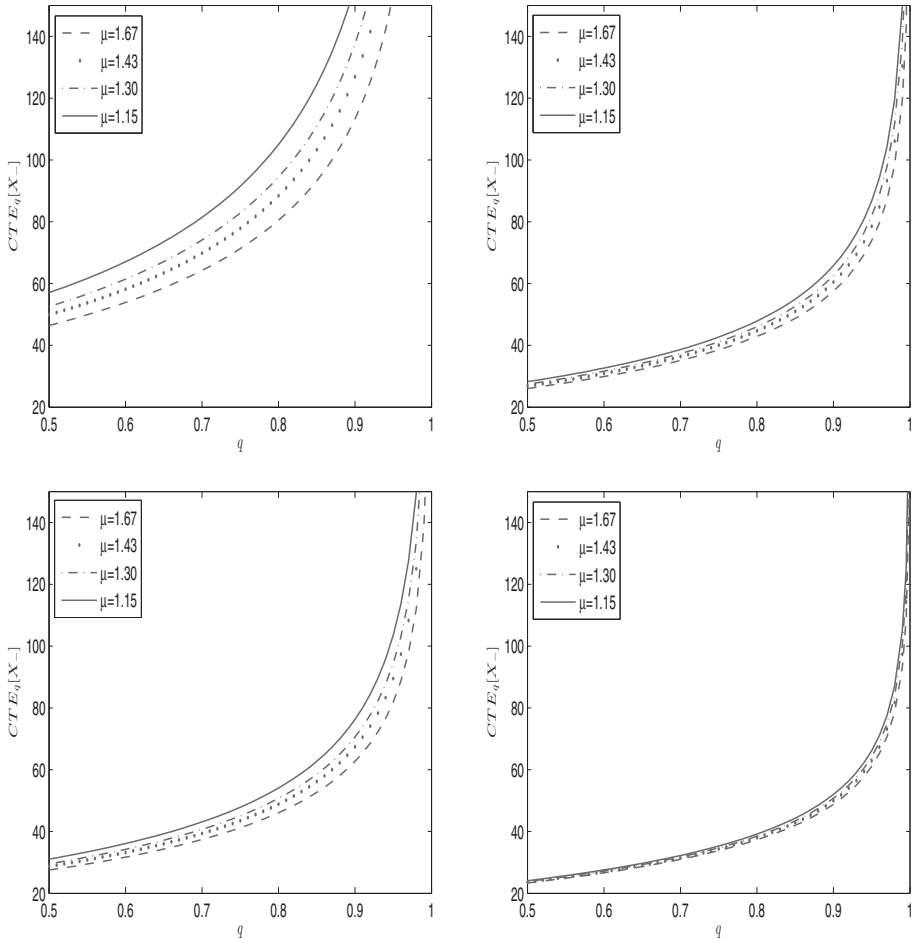


FIGURE 2: Conditional expected times of first default for portfolios (1) (top left panel), (2) (top right panel), (3) (bottom left panel) and reference (\perp) (bottom right panel) with the parameter μ varying from 1.67 to 1.15 and the default probability $p = 0.3198$. Proposition 4.2 is employed to compute the values of CTE_q for $q \in [0, 1)$. (Color online)

2.5 along with (5.1) results in

$$\overline{F}_+^{(\perp)} \geq_{st} \overline{F}_+^{(2)} \geq_{st} \overline{F}_+^{(3)} \geq_{st} \overline{F}_+^{(1)},$$

and hence

$$CTE_q^{(\perp)}[X_+] \geq CTE_q^{(2)}[X_+] \geq CTE_q^{(3)}[X_+] \geq CTE_q^{(1)}[X_+],$$

for all $q \in [0, 1)$ and $X_+ \in \mathcal{X}$. This conforms with the right panel of Figure 1.

Unlike in the case of the first default, we observe that if the time of the last default is of interest and the distributions of the r.c.'s are fixed, then assuming

stronger correlations between r.c.'s yields a more conservative assessment of the expected time of the last default.

6. CONCLUSIONS

We have introduced and studied a new form of an absolutely continuous with respect to the Lebesgue measure multivariate probability law with the univariate margins distributed Pareto of the second kind. The genesis of our distribution is threefold, i.e., it originates as the Laplace transform of a multivariate gamma distribution with the dependence structure based on the additive form of the multivariate reduction method, and it also admits variants of the multiplicative background risk model as well as the minima-based common shock model. We have meaningfully positioned the proposed multivariate Pareto distribution in the general context of the current state of the art. We have proved and employed certain characteristic results to derive, e.g., the (conditional/product) moments of the new multivariate Pareto as well as the distributions of minima and maxima. Last but not least, we have developed expressions for some tail-based risk measures of actuarial interest and elucidated our findings with the help of a numerical example.

ACKNOWLEDGEMENTS

We thank the anonymous referees and the Associate Editor, Prof. Montserrat Guillén, for valuable comments and suggestions that improved the work significantly and resulted in a better presentation of the material. We are also grateful to Prof. Paul Embrechts and all participants of the ETHs Series of Talks in Financial and Insurance Mathematics for feedback and insights.

Our research has been supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. Jianxi Su also acknowledges the financial support of the Government of Ontario and MITACS Canada via, respectively, the Ontario Graduate Scholarship program and the Elevate Postdoctoral fellowship.

REFERENCES

- ARNOLD, B.C. (1983) *Pareto Distributions*. Fairland: International Cooperative Publishing House.
- ASIMIT, V., FURMAN, E. and VERNIC, R. (2010) On a multivariate Pareto distribution. *Insurance: Mathematics and Economics*, **46**(2), 308–316.
- ASIMIT, V., VERNIC, R. and ZITIKIS, R. (2013) Evaluating risk measures and capital allocations based on multi-losses driven by a heavy-tailed background risk: The multivariate Pareto-II model. *Risk*, **1**(1), 14–33.
- ASIMIT, V., VERNIC, R. and ZITIKIS, R. (2016) Background risk models and stepwise portfolio construction. *Methodology and Computing in Applied Probability*, in press.
- BALKEMA, A. and DE HAAN, L. (1974) Residual life time at great age. *Annals of Probability*, **2**(5), 792–804.

- BENSON, D.A., SCHUMER, R. and MEERSCHAERT, M.M. (2007) Recurrence of extreme events with power-law interarrival times. *Geophysical Research Letters*, **34**(16), 1–5.
- BOUCHER, J.P., DENUIT, M. and GUILLÉN, M. (2008) Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions. *Variance*, **2**(1), 135–162.
- BOWERS, N.L., GERBER, H.U., HICKMAN, J.C., JONES, D.A. and NESBITT, C.J. (1997) *Actuarial Mathematics*, 2nd ed. Schaumburg: Society of Actuaries.
- CEBRIÁN, A.C., DENUIT, M. and LAMBERT, P. (2003) Generalized Pareto fit to the society of actuaries large claims database. *North American Actuarial Journal*, **7**(3), 18–36.
- CHAVEZ-DEMOULIN, V., EMBRECHTS, P. and HOFERT, M. (2015) An extreme value approach for modeling operational risk losses depending on covariates. *Journal of Risk and Insurance*, in press.
- CHERIAN, K.C. (1941) A bivariate correlated gamma-type distribution function. *Journal of the Indian Mathematical Society* **5**, 133–144.
- CHIRAGIEV, A. and LANDSMAN, Z. (2009) Multivariate flexible Pareto model: Dependency structure, properties and characterizations. *Statistics and Probability Letters*, **79**(16), 1733–1743.
- CHOO, W. and DE JONG, P. (2009) Loss reserving using loss aversion functions. *Insurance: Mathematics and Economics*, **45**(2), 271–277.
- CHOO, W. and DE JONG, P. (2010) Determining and allocating diversification benefits for a portfolio of risks. *ASTIN Bulletin*, **40**(1), 257–269.
- EMBRECHTS, P., MCNEIL, A. and STRAUMANN, D. (2002) Correlation and dependence in risk management: Properties and pitfalls. In *Risk Management: Value at Risk and Beyond* (ed. M. Dempster), Cambridge: Cambridge University Press.
- ENGELMANN, B. and RAUHMEIER, R. (2011) *The Basel II Risk Parameters: Estimation, Validation, Stress Testing - with Applications to Loan Risk Management*. Berlin: Springer.
- FELLER, W. (1966) *An Introduction to Probability Theory and its Applications*. New York: John Wiley and Sons.
- FRANKE, G., SCHLESINGER, H. AND STAPLETON, R.C. (2006) Multiplicative background risk. *Management Science*, **52**(1), 146–153.
- FURMAN, E. (2008) On a multivariate gamma distribution. *Statistics and Probability Letters*, **78**(15), 2353–2360.
- FURMAN, E. and LANDSMAN, Z. (2005) Risk capital decomposition for a multivariate dependent gamma portfolio. *Insurance: Mathematics and Economics*, **37**(3), 635–649.
- FURMAN, E. and LANDSMAN, Z. (2010) Multivariate Tweedie distributions and some related capital-at-risk analysis. *Insurance: Mathematics and Economics*, **46**(2), 351–361.
- FURMAN, E. and ZITIKIS, R. (2008a) Weighted premium calculation principles. *Insurance: Mathematics and Economics*, **42**(1), 459–465.
- FURMAN, E. and ZITIKIS, R. (2008b) Weighted risk capital allocations. *Insurance: Mathematics and Economics* **43**(2), 263–269.
- FURMAN, E. and ZITIKIS, R. (2010) General Stein-type covariance decompositions with applications to insurance and finance. *ASTIN Bulletin* **40**(1), 369–375.
- GABAIX, X., GOPIKRISHNAN, P., PLEROU, V. and STANLEY, H.E. (2003) A theory of power-law distributions in financial market fluctuations. *Nature*, **423**, 267–270.
- GOLLIER, C. and PRATT, J.W. (1996) Weak proper risk aversion and the tempering effect of background risk. *Econometrica* **64**(5), 1109–1123.
- GRADSHTEYN, I.S. and RYZHIK, I.M. (2007) *Table of Integrals, Series and Products*. 7th ed. New York: Academic Press.
- KOEDIJK, K.G., SCHAFGANS, M.M.A. and DE VRIES, C.G. (1990) The tail index of exchange rate returns. *Journal of International Economics*, **29**(1–2), 93–108.
- KOTZ, S., BALAKRISHNAN, N. and JOHNSON, N.L. (2000) *Continuous Multivariate Distributions*. 2nd ed. New York: Wiley.
- LONGIN, F.M. (1996) The asymptotic distribution of extreme stock market returns. *Journal of Business*, **69**(3), 383–408.
- MATHAI, A.M. and MOSCHOPOULOS, P.G. (1991) On a multivariate gamma. *Journal of Multivariate Analysis* **39**(1), 135–153.
- MATHAI, A.M. and MOSCHOPOULOS, P.G. (1992) A form of multivariate gamma distribution. *Annals of the Institute of Statistical Mathematics*, **44**(1), 97–106.

- MEYERS, G.G. (2007) The common-shock model for correlated insurance losses. *Variance*, **1**(1), 40–52.
- MOSCHOPOULOS, P.G. (1985) The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics*, **37**, 541–544.
- PARETO, V. (1897) The new theories of economics. *Journal of Political Economy*, **5**(4), 485–502.
- PFEIFER, D. and NEŠLEHOVÁ, J. (2004) Modeling and generating dependent risk processes for IRM and DFA. *ASTIN Bulletin*, **34**(2), 333–360.
- PICKANDS, J. (1975) Statistical inference using extreme order statistics. *The Annals of Statistics*, **3**(1), 119–131.
- RAMABHADRAN, V. (1951) A multivariate gamma-type distribution. *Journal of Multivariate Analysis*, **38**, 213–232.
- SOPRANO, A., CRIELAARD, B., PIACENZA, F. and RUSPANTINI, D. (2009) *Measuring Operational and Reputational Risk: A Practitioner's Approach*. Chichester: Wiley.
- STANDARD & POOR'S (2015) Default, transition and recovery: 2014 annual global corporate default study and rating transitions. Technical report, Standard and Poor's, New York.
- TSANAKAS, A. (2008) Risk measurement in the presence of background risk. *Insurance: Mathematics and Economics*, **42**(2), 520–528.
- VERNIC, R. (1997) On the bivariate generalized Poisson distribution. *ASTIN Bulletin*, **27**(1), 23–32.
- VERNIC, R. (2000) A multivariate generalization of the generalized Poisson distribution. *ASTIN Bulletin*, **30**(1), 57–67.
- VERNIC, R. (2011) Tail conditional expectation for the multivariate Pareto distribution of the second kind: Another approach. *Methodology and Computing in Applied Probability*, **13**(1), 121–137.
- WANG, S. (1996) Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, **26**(1), 71–92.

JIANXI SU

Department of Mathematics and Statistics
York University
Toronto, Ontario M3J 1P3
Canada

EDWARD FURMAN (Corresponding author)

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

APPENDIX A. PROOFS

Proof of Proposition 2.1. By construction, we readily have that

$$\hat{G}_{1,\dots,n}(\mathbf{t}) = \mathbf{E} \left[e^{-\mathbf{X}'\mathbf{t}} \right] = \mathbf{E} \left[\exp \left\{ - \sum_{i=1}^n \sum_{j=1}^{n+1} \frac{c_{i,j}}{\sigma_i} Y_j t_i \right\} \right] = \prod_{j=1}^{n+1} \hat{G}_j \left(\sum_{i=1}^n \frac{c_{i,j}}{\sigma_i} t_i \right),$$

which along with (1.2) completes the proof. ■

Proof of Theorem 2.1. Let G^Π denote a multivariate c.d.f. with mutually independent gamma-distributed univariate margins. We have the following string of equations:

$$\begin{aligned} \left(\prod_{i=1}^n \sigma_i\right) f_{1,\dots,n}(x_1, \dots, x_n) &= \left(\prod_{i=1}^n \sigma_i\right) (-1)^n \frac{\partial^n}{\partial x_1 \dots \partial x_n} \bar{F}_{1,\dots,n}(x_1, \dots, x_n) \\ &= \mathbf{E} \left[\left(\prod_{i=1}^n \sigma_i\right) (-1)^n \frac{\partial^n}{\partial x_1 \dots \partial x_n} \exp \left\{ -\sum_{i=1}^n \sum_{j=1}^{n+1} \frac{c_{i,j}}{\sigma_i} Y_j x_i \right\} \right] \\ &= \mathbf{E} \left[\left(\prod_{i=1}^n \sum_{j=1}^{n+1} c_{i,j} Y_j\right) \exp \left\{ -\sum_{i=1}^n \sum_{j=1}^{n+1} \frac{c_{i,j}}{\sigma_i} Y_j x_i \right\} \right] \\ &= \sum_{\forall i_j \in I} d_c(i_1, \dots, i_{n+1}) \int_{\mathbf{R}_+^{n+1}} \prod_{j=1}^{n+1} \exp \left\{ -y_j \left(\sum_{i=1}^n c_{i,j} \frac{x_i}{\sigma_i} \right) \right\} \\ &\quad \times \frac{\Gamma(\gamma_j + i_j)}{\Gamma(\gamma_j)} dG_{1,\dots,n+1}^\Pi(\mathbf{y}; \tilde{\boldsymbol{\gamma}}, 1), \end{aligned}$$

where $\tilde{\boldsymbol{\gamma}} = (\gamma_1 + i_1, \dots, \gamma_{n+1} + i_{n+1})'$ is a vector of positive parameters. The proof is completed by computing the iterated integral. ■

Proof of Theorem 2.2. Let F_{Ξ} denote the c.d.f. of the r.v. $\Xi = (\Xi_1, \dots, \Xi_n)'$. The “if” part is immediate from the following obvious relations:

$$\begin{aligned} \bar{F}_{1,\dots,n}(x_1, \dots, x_n) &= \mathbf{P}[\Lambda_1 > \Xi_1 x_1, \dots, \Lambda_n > \Xi_n x_n] \\ &= \int_{\mathbf{R}_+^n} \mathbf{P}[\Lambda_1 > \xi_1 x_1, \dots, \Lambda_n > \xi_n x_n] dF_{\Xi}(\xi_1, \dots, \xi_n) \\ &= \int_{\mathbf{R}_+^n} \exp \left\{ -\sum_{i=1}^n \xi_i x_i \right\} dF_{\Xi}(\xi_1, \dots, \xi_n), \end{aligned} \tag{A.1}$$

and by Proposition 2.1. The “only if” part follows because (A.1) is the n -variate Laplace transform of $Ga_{1,\dots,n}(\boldsymbol{\gamma}_c^*, \boldsymbol{\sigma})$, and it is thus unique. This completes the proof. ■

Proof of Theorem 2.3. For the proof, we readily have that

$$\begin{aligned} \bar{F}_-(x) &= \bar{F}_{1,\dots,n}(x, \dots, x) = \prod_{j=1}^{n+1} \left(1 + \sum_{i=1}^n \frac{c_{i,j}}{\sigma_i} x \right)^{-\gamma_j} \\ &= \int_0^\infty e^{-\lambda x} dF_{Z_1+\dots+Z_{n+1}}(\lambda), \text{ where } x \in \mathbf{R}_+, \end{aligned}$$

which establishes the mixture representation. ■

Proof of Theorem 2.4. Employing Theorem 2.3 with $Z_j \sim Ga(\gamma_j, (\sum_{i=1}^n \frac{c_{i,j}}{\sigma_i})^{-1})$, $j = 1, \dots, n + 1$, Lemma 2.1, changing the order of summation and integration and using

equation (1.2), we have that

$$\begin{aligned} \bar{F}_-(x) &= \int_0^\infty \sum_{k=0}^\infty e^{-\lambda x} p_k \frac{e^{-\lambda \sigma + \lambda \gamma^* + k - 1} \alpha_+(\boldsymbol{\sigma})^{\gamma^* + k}}{\Gamma(\gamma^* + k)} d\lambda \\ &= \sum_{k=0}^\infty \left(1 + \frac{x}{\alpha_+(\boldsymbol{\sigma})}\right)^{-(\gamma^* + k)} p_k. \end{aligned}$$

This completes the proof. ■

Proof of Theorem 3.1. Let G^Π denote a multivariate c.d.f. with mutually independent gamma-distributed univariate margins. We start by employing Lemma 2.2 and observation (3.1), and then do change of variables and obtain that

$$\begin{aligned} (\sigma_k \sigma_l)^{-1} \mathbf{E}[X_k X_l] &= (\sigma_k \sigma_l)^{-1} \mathbf{E} \left[\frac{1}{\Xi_k} \cdot \frac{1}{\Xi_l} \right] \\ &= \int_{\mathbf{R}_+^3} \frac{1}{(y_3 + y_1)(y_3 + y_2)} dG_{1,\dots,3}^\Pi(\mathbf{y}; \boldsymbol{\gamma}, \mathbf{1}) \\ &= \int_0^\infty \int_0^\infty (1+v)^{-\gamma_{c,l}} (1+u)^{-\gamma_{c,k}} (1+u+v)^{-\gamma_{c,(k,l)}} dudv \\ &= \int_0^\infty (1+v)^{-\gamma_{c,l}^*} \left(\int_0^\infty (1+u)^{-\gamma_{c,k}} \left(1 + \frac{u}{1+v}\right)^{-\gamma_{c,(k,l)}} du \right) dv \\ &\stackrel{(1)}{=} \int_0^\infty (1+v)^{-\gamma_{c,l}^*} \frac{1}{\gamma_{c,k}^* - 1} {}_2F_1 \left(\gamma_{c,(k,l)}, 1; \gamma_{c,k}^*; \frac{v}{1+v} \right) dv \\ &= \int_0^1 (1-u)^{(\gamma_{c,l}^* - 1) - 1} \frac{1}{\gamma_{c,k}^* - 1} {}_2F_1(\gamma_{c,(k,l)}, 1; \gamma_{c,k}^*; u) du \\ &\stackrel{(2)}{=} \frac{1}{(\gamma_{c,k}^* - 1)(\gamma_{c,l}^* - 1)} {}_3F_2(\gamma_{c,(k,l)}, 1, 1; \gamma_{c,k}^*, \gamma_{c,l}^*; 1), \end{aligned}$$

where “ $\stackrel{(1)}{=}$ ” holds because of the following integral representation of the Gauss hypergeometric function (Equation 3.197(5) in Gradshteyn and Ryzhik, 2007)

$${}_2F_1(\alpha, \beta; \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma - \beta)} \int_0^\infty t^{\beta-1} (1+t)^{\alpha-\gamma} (1+zt)^{-\alpha} dt,$$

for $\gamma > \beta > 0$ and all z such that the integral above converges, and “ $\stackrel{(2)}{=}$ ” follows from Equation 7.512(5) in loc. cit. This completes the proof. ■

Proof of Corollary 3.1. The lower bound follows by setting $\gamma_{c,(k,l)} \equiv 0$ and both of $\gamma_{c,l}^*$ and $\gamma_{c,k}^*$ to exceed 2. To establish the upper bound, let $\gamma_{c,k} \rightarrow 0$ and $\gamma_{c,l} \rightarrow 0$ and assume that

$\gamma_{c,(k,l)}$ exceeds 2, and then we have that

$$\begin{aligned} &\lim_{\gamma_{c,k}, \gamma_{c,l} \rightarrow 0} {}_3F_2(\gamma_{c,(k,l)}, 1, 1; \gamma_{c,k} + \gamma_{c,(k,l)}, \gamma_{c,l} + \gamma_{c,(k,l)}; 1) \\ &= {}_3F_2(\gamma_{c,(k,l)}, 1, 1; \gamma_{c,(k,l)}, \gamma_{c,(k,l)}; 1) \\ &= {}_2F_1(1, 1; \gamma_{c,(k,l)}; 1) \\ &= \frac{\gamma_{c,(k,l)} - 1}{\gamma_{c,(k,l)} - 2}, \end{aligned}$$

where the last equality holds due to Equation 9.122(1) in Gradshteyn and Ryzhik (2007), and the covariance of interest reduces to

$$\text{Cov}[X_k, X_l] \rightarrow \frac{\sigma_l \sigma_k}{(\gamma_{c,(k,l)} - 1)^2 (\gamma_{c,(k,l)} - 2)}, \text{ where } 1 \leq k \neq l \leq n.$$

This, along with Proposition 2.2, completes the proof. ■

Proof of Corollary 3.2. To obtain (3.3), we set $c_{i,i} = c_{i,n+1} \equiv 1$ and zero otherwise. This implies $\gamma_{c,k}^* = \gamma_k + \gamma_{n+1}$, $\gamma_{c,l}^* = \gamma_l + \gamma_{n+1}$ and $\gamma_{c,(k,l)} = \gamma_{n+1}$ (see, Example 1.1). Then the result directly follows from Theorem 3.1. To establish (3.4), let $c_{i,j} \equiv 1$ for $1 \leq j \leq i \leq n$ and zero otherwise. Then the desired assertion follows because

$$\begin{aligned} \text{Cov}[X_{2,k}, X_{2,l}] &= \frac{\sigma_k \sigma_l}{(\gamma_{c,k}^* - 1)(\gamma_{c,l}^* - 1)} ({}_3F_2(\gamma_{c,(k,l)}, 1, 1; \gamma_{c,(k,l)}, \gamma_{c,l}^*; 1) - 1) \\ &= \frac{\sigma_k \sigma_l}{(\gamma_{c,k}^* - 1)(\gamma_{c,l}^* - 1)} ({}_2F_1(1, 1; \gamma_{c,l}^*; 1) - 1) \\ &= \frac{\sigma_k \sigma_l}{(\gamma_{c,k}^* - 1)(\gamma_{c,l}^* - 1)(\gamma_{c,l}^* - 2)}, \end{aligned}$$

where the latter equality holds for $\gamma_{c,l}^* > 2$. This completes the proof of the corollary. ■

Proof of Theorem 3.2. We first note that

$$\mathbf{P}[X_k > x_k | X_l = x_l] = \frac{-\frac{\partial}{\partial x_l} \bar{F}_{X_k, X_l}(x_k, x_l)}{f_{X_l}(x_l)} \text{ for } x_k \text{ and } x_l \text{ both in } \mathbf{R}_+,$$

and then write

$$\begin{aligned} -\frac{\partial}{\partial x_l} \bar{F}_{X_k, X_l}(x_k, x_l) &= \left(1 + \frac{x_k}{\sigma_k}\right)^{-\gamma_{c,k}} \left[\frac{\gamma_{c,(k,l)}}{\sigma_l} \left(1 + \frac{x_k}{\sigma_k} + \frac{x_l}{\sigma_l}\right)^{-\gamma_{c,(k,l)}-1} \left(1 + \frac{x_l}{\sigma_l}\right)^{-\gamma_{c,l}} \right. \\ &\quad \left. + \frac{\gamma_{c,l}}{\sigma_l} \left(1 + \frac{x_k}{\sigma_k} + \frac{x_l}{\sigma_l}\right)^{-\gamma_{c,(k,l)}} \left(1 + \frac{x_l}{\sigma_l}\right)^{-\gamma_{c,l}-1} \right]. \end{aligned} \tag{A.2}$$

Plain simplifications complete the proof. ■

Proof of Theorem 3.3. Note that for the pair $(X_k, X_l)'$, $0 \leq k \neq l \leq n$, stochastic representation (3.1) is of utmost generality, i.e., conditional distribution function (4.10) coincides with one of the type I multivariate flexible Pareto model. Hence, the proof is completed by evoking Theorem 3 of Chiragiev and Landsman (2009) as well as Proposition 2.2. ■

Proof of Theorem 4.1. Note that

$$\wedge_{j=1}^{n+1} (X_{\lambda_j} * \Lambda_j) \stackrel{d}{=} \left(\wedge_{j=1}^{n+1} X_{\lambda_j} \right) * \Lambda,$$

and hence, for F_Λ denoting the c.d.f. of the r.v. $\Lambda = (\Lambda_1, \dots, \Lambda_{n+1})'$,

$$\bar{F}_{1,\dots,n}(x_1, \dots, x_n) = \int_{\mathbf{R}_+^{n+1}} \prod_{i=1}^n \exp \left\{ - \sum_{j=1}^{n+1} \frac{c_{i,j} x_i \lambda_j}{\sigma_i} \right\} dF_\Lambda(\lambda_1, \dots, \lambda_{n+1}),$$

which, along with (1.2), completes the proof. ■

Proof of Proposition 4.1. Notice that

$$\pi_w[X] = \frac{\mathbf{E}[\mathbf{E}[w(X) | \Lambda] \pi_w[X | \Lambda]]}{\mathbf{E}[\mathbf{E}[w(X) | \Lambda]]} \quad \text{for } X \in \mathcal{X},$$

and set $w_1(\lambda) = \mathbf{E}[w(X) | \lambda]$ and $v_1(\lambda) = \pi_w[X | \lambda]$. This concludes the proof. ■

Proof of Corollary 4.1. Use Proposition 4.1 setting the weight function equal to the Dirac delta function, or alternatively evoke Proposition 2.2. This concludes the proof. ■

Proof of Corollary 4.2. Fix $w(x) = \mathbf{1}\{x > VaR_q[X]\}$ for $q \in [0, 1)$, the result follows from Proposition 4.1. ■

Proof of Corollary 4.3. As $X_i | \Lambda = \lambda \sim Exp(\lambda)$, $\lambda > 0$, $i = 1, \dots, n$, we readily have that

$$CTE_{q^*}[X_i | \Lambda = \lambda] = \frac{1}{\lambda} + VaR_q[X_i],$$

and

$$\bar{C}[VaR_q[X_i]; \lambda] = e^{-\lambda VaR_q[X_i]},$$

and the assertion holds by Proposition 4.1. This completes the proof. ■

Proof of Proposition 4.2. We readily have that

$$\begin{aligned} CTE_q[X_-] &= VaR_q[X_-] + \frac{1}{1-q} \int_{VaR_q[X_-]}^\infty \bar{F}_-(x) dx \\ &\stackrel{(1)}{=} VaR_q[X_-] + \frac{1}{1-q} \int_{VaR_q[X_-]}^\infty \left(\sum_{k=0}^\infty \left(1 + \frac{x}{\alpha_+(\sigma)} \right)^{-(\gamma^*+k)} p_k \right) dx \\ &\stackrel{(2)}{=} VaR_q[X_-] + \frac{\mathbf{E}[X_-]}{1-q} \int_{VaR_q[X_-]}^\infty \\ &\quad \times \left(\sum_{k=0}^\infty \frac{\gamma^* + k - 1}{\alpha_+(\sigma)} \left(1 + \frac{x}{\alpha_+(\sigma)} \right)^{-(\gamma^*+k-1)} q_k \right) dx, \end{aligned}$$

where “ $\stackrel{(1)}{=}$ ” follows because of Corollary 2.4 and “ $\stackrel{(2)}{=}$ ” holds since

$$\mathbf{E}[X_-] = \sum_{k=0}^\infty \frac{\alpha_+(\sigma)}{\gamma^* + k - 1} p_k.$$

This completes the proof. ■

Proof of Proposition 4.3. We observe that

$$\int_x^\infty \bar{F}(t)dt = \mathbf{E}[X - x | X > x]\bar{F}(x),$$

for all x in the range of the r.v. X . Then the assertion follows by Proposition 2.5 and changing the order of summation and integration in

$$CTE_q[X_+] = \frac{\int_{VaR_q[X_+]}^\infty \sum_{S \subseteq \{1, \dots, n\}} (-1)^{|S|-1} \bar{F}_{S^c}(x) dx}{1 - q} + VaR_q[X_+].$$

This completes the proof. ■

Proof of Proposition 4.5. Note that, for $1 \leq k \neq l \leq n$,

$$\begin{aligned} \mathbf{E}[X_k | X_l > VaR_q[X_l]] &= \int_0^\infty \mathbf{P}[X_k > x | X_l > VaR_q[X_l]] dx \\ &\stackrel{(1)}{=} \sigma_k \int_0^\infty (1 + u)^{-\gamma_{c,k}} \left(1 + \frac{u}{1 + VaR_q[X_l]/\sigma_l} \right)^{-\gamma_{c,(k,l)}} du, \end{aligned}$$

where “ $\stackrel{(1)}{=}$ ” holds because of observation (3.1) and techniques similar to the ones used in Theorem 3.1. The Proposition then follows by evoking Equation 3.197(5) in Gradshteyn and Ryzhik (2007). This completes the proof. ■