ON COMPOSITE PARETO MODELS

SANDRA TEODORESCU and RALUCA VERNIC

Communicated by the former editorial board

To model statistical data coming from two different distributions, Cooray and Ananda [1] introduced a composite Lognormal-Pareto model, that was further developed by Scollnik [6]. In this paper, we consider a more general composite Pareto model, obtained by replacing the Lognormal distribution by an arbitrary continuous distribution. The main characteristics of this model, as well as some statistical inference are presented. The particular cases of the composite Gamma-Pareto and Weibull-Pareto models are detailed and numerically illustrated.

AMS 2010 Subject Classification: 60E05, 62F10, 46N30, 62P05.

Key words: composite distributions, Pareto distribution, Gamma distribution, Weibull distribution, statistical inference.

1. INTRODUCTION

In some situations, statisticians encounter data that obviously come from two different models. This is often the case with e.g. insurance payments data, when actuaries must handle smaller data with high frequencies and occasional larger data with lower frequencies (see, e.g., [3], [4]). Then, the corresponding distribution can be modeled as a combination of two distributions, consisting of a less heavy-tailed distribution up to a certain threshold value, and a heavy-tailed distribution from that threshold on. Such distributions are the composite ones as suggested by Cooray and Ananda [1]. They constructed a composite model having the probability density function

$$f(x) = \begin{cases} cf_1(x), & -\infty < x \leq \theta, \\ cf_2(x), & \theta < x < \infty, \end{cases} \tag{1.1}$$

where f_1 and f_2 are probability density functions, while c is a normalizing constant that results by imposing continuity and differentiability conditions at θ. In [4], (1.1) is called a two-component spliced model and it is recommended when the tail behavior is inconsistent with the small losses behavior. Therefore, the role of the threshold θ is quite important, being the point where, based on data analysis, we change the model; it is chosen between 0 and the contract's upper limit. For example, we sometimes notice a value such that

there is a large amount of data below that value, and a limited amount of information above it. This makes the probability of losses above θ clearly smaller than the probability of losses below it.

In Cooray and Ananda [1], f_1 is taken to be the Lognormal density, while f_2 is the Pareto one. Based on the same distributions, Scollnik [6] proposed two different composite models, which are more general than the one studied by Cooray and Ananda [1]. In the same manner, Teodorescu and Vernic [7] suggested composite Exponential-Pareto models. Moreover, Preda and Ciumara [5] realized a comparative study between the Weibull-Pareto and the Lognormal-Pareto composite distributions.

In this paper, we go even further, by considering a more general composite Pareto model, obtained by letting f_1 be an arbitrary probability density function, while f_2 remains the Pareto density function. The choice of the Pareto distribution is not incidentally: it is the classical heavy-tailed distribution, most used e.g. to model actuarial heavy-tailed claims. If we aim to model insurance claims, then, the choice of f_1 will be restricted to distributions less-heavy tailed than Pareto and defined only for positive values.

Section 2 is dedicated to the study of general composite models. We start Subsection 2.1 by recalling a mixture model equivalent to model (1.1), with accent on its density function, cumulative distribution function (cdf), initial moments, characteristic function, and two methods for estimating the parameters. Subsection 2.2 is dedicated to a similar study of the composite Pareto model, obtained from the general model studied in Subsection 2.1 as described before.

In Section 3, we introduce two new composite Gamma-Pareto models while in Section 4 we study two new composite Weibull-Pareto models. We discuss some of their properties based on the general theory from Section 2. The suggested estimation methods are illustrated on generated data. We end with a conclusions section.

In the following, we denote by $\mathbb{R}$ the set of real numbers and by $\mathbb{N}$ the set of all positive integers. Also, if f is a real function, then f' denotes its first derivative, while f'' its second derivative.

2. COMPOSITE MODELS

2.1. SOME PROPERTIES OF A MIXTURE MODEL

2.1.1. **The model.** Scollnik [6] noticed that the density function of the composite model (1.1) could also be written as

$$f(x) = \begin{cases} r f_1^*(x), & -\infty < x \leq \theta, \\ (1-r) f_2^*(x), & \theta < x < \infty, \end{cases} \tag{2.1}$$

where $0 \leq r \leq 1$, while f_1^* and f_2^* are adequate truncations of the density functions f_1 and f_2. More precisely, if F_i denotes the cdf of f_i, we have

$$
\begin{aligned}
f_1^*(x) &= \frac{f_1(x)}{F_1(\theta)}, \qquad -\infty < x \leq \theta, \\
f_2^*(x) &= \frac{f_2(x)}{1 - F_2(\theta)}, \quad \theta < x < \infty.
\end{aligned} \tag{2.2}
$$

It is easy to see that the density function (2.1) can be interpreted as a two-component mixture model with mixing weights r and $1 - r$ (i.e., a convex combination of two density functions),

$$
f(x) = r f_1^*(x) + (1 - r)\, f_2^*(x), \quad r \in [0, 1]. \tag{2.3}
$$

In general, we would like the density function (2.1) to be smooth, so we impose continuity and differentiability conditions at θ. Assuming that all the quantities involved in the following exist, it is easy to see that the continuity condition $f(\theta - 0) = f(\theta + 0)$, yields $r = \frac{f_2(\theta)F_1(\theta)}{f_2(\theta)F_1(\theta)+f_1(\theta)(1-F_2(\theta))}$ while the differentiability condition at θ, $f'(\theta - 0) = f'(\theta + 0)$, yields $r = \frac{f_2'(\theta)F_1(\theta)}{f_2'(\theta)F_1(\theta)+f_1'(\theta)(1-F_2(\theta))}$. If we combine the two results for r, we also obtain a restriction for θ, i.e., $\frac{f_1(\theta)}{f_2(\theta)} = \frac{f_1'(\theta)}{f_2'(\theta)}$.

In [7], Teodorescu and Vernic deduced some properties of model (2.1) concerning its cdf, initial moments and characteristic function. These properties are presented below. If X is a random variable (r.v.) with density function f, we denote its n-th order initial moment by $E_n(f) = E(X^n)$, and its characteristic function by $\varphi_f(t) = \varphi_X(t) = E\left(e^{itX}\right)$.

PROPOSITION 2.1. a) *Let F denote the cdf of the density function given in* (2.1). *Then*

$$
F(x) = \begin{cases} r\dfrac{F_1(x)}{F_1(\theta)}, & -\infty < x \leq \theta, \\ r + (1 - r)\dfrac{F_2(x) - F_2(\theta)}{1 - F_2(\theta)}, & \theta < x < \infty. \end{cases}
$$

b) *Assuming that all the quantities involved exist, the n-th order initial moment of the density function* (2.1) *is*

$$
E_n(f) = r E_n(f_1^*) + (1 - r)\, E_n(f_2^*),
$$

while its characteristic function is

$$
\varphi_f(t) = r\varphi_{f_1^*}(t) + (1 - r)\varphi_{f_2^*}(t), \quad t \in \mathbb{R}.
$$

2.1.2. **Generating random values from the composite density function (2.1).** In order to generate such random values, we suggest using

the inversion method, assuming that both F_1 and F_2 admit inverse functions. Therefore, if u is a value generated from the uniform distribution U(0,1), we obtain a value x generated from (2.1) as

If $u \leq F(\theta) = r$ then solve $u = r\frac{F_1(x)}{F_1(\theta)}$ for x, i.e., $x = F_1^{-1}\left(\frac{u}{r}F_1(\theta)\right)$.

If $u > r$ then solve $u = r + (1-r)\frac{F_2(x)-F_2(\theta)}{1-F_2(\theta)}$ for x, i.e.,

$$x = F_2^{-1}\left(\frac{u - r + (1-u)F_2(\theta)}{1-r}\right).$$

2.1.3. **Statistical inference**

I. A first algorithm. In Teodorescu and Vernic [7], the following algorithm based on the maximum likelihood (ML) method was suggested. Assume that the density function (2.1) depends on the real parameters $\delta_1, \ldots, \delta_s, \theta$, where $s \in \mathbb{N}$, and consider the random data sample $(x_1, \ldots, x_n)$. Without loss of generality, we assume that it is an ordered sample, i.e., $x_1 \leq x_2 \leq \cdots \leq x_n$. In order to apply the ML method, we must know the integer value m such that the unknown parameter θ is in between the m-th and $m+1$-th observations, i.e., $x_m \leq \theta \leq x_{m+1}$. Assuming that somehow we know this m, the likelihood function is

$$L(x_1, \ldots, x_n; \delta_1, \ldots, \delta_s, \theta) = \prod_{i=1}^{m} r f_1^*(x_i) \prod_{j=m+1}^{n} (1-r) f_2^*(x_j) = \tag{2.4}$$

$$= r^m (1-r)^{n-m} \prod_{i=1}^{m} f_1^*(x_i) \prod_{j=m+1}^{n} f_2^*(x_j).$$

Unfortunately, in general, we don't know the exact value of m; also, notice that if m changes, the ML estimation also changes. Therefore, we suggest the following estimation algorithm that takes into consideration all possible values of m such that $x_m \leq \theta \leq x_{m+1}$:

Step 1. For each $m = 1, 2, \ldots, n-1$, do

Evaluate $\hat{\delta}_1, \ldots, \hat{\delta}_s, \hat{\theta}$ as solutions of the ML system

$$\begin{cases} \dfrac{\partial \ln L}{\partial \delta_i} = 0, \quad i = 1, \ldots, s, \\ \dfrac{\partial \ln L}{\partial \theta} = 0. \end{cases} \tag{2.5}$$

If $x_m \leq \hat{\theta} \leq x_{m+1}$ then the ML estimations are

$$\hat{\delta}_i^{ML} = \hat{\delta}_i, \quad i = 1, \ldots, s, \quad \hat{\theta}^{ML} = \hat{\theta}.$$

Step 2. If Step 1 doesn't give any solution for θ, then we are in one of two situations: $m = n$ or $m = 0$, hence, we recommend using only f_1 and, respectively, f_2 for the likelihood function.

Remark 2.1. With this algorithm, one has to check $n-1$ intervals, so that the computing time strongly depends on the magnitude of n. Moreover, if n is large and (2.5) leads to a very complex system that needs numerical evaluation, this algorithm might be very difficult to implement. Therefore, we suggest an alternative algorithm based on the method of moments and on quantiles matching, that can at least provide starting values for the above algorithm.

II. A second algorithm. As already seen in the first algorithm, our main problem is that we don't know anything about the magnitude of the parameter θ. Therefore, assuming as before that (2.1) depends on $s+1$ unknown parameters, we suggest the following alternative algorithm:

Step 1. Let q_1 and q_3 be the first and, respectively, third empirical quartiles of the data sample. We first assume that $q_1 < \theta < q_3$. Then, we use the method of moments to match the first $s-1$ empirical moments with the corresponding theoretical ones, and we add two more equations from matching the two quartiles, i.e., add

$$\begin{cases} r\dfrac{F_1(q_1)}{F_1(\theta)} = 0.25, \\ r + (1-r)\dfrac{F_2(q_3) - F_2(\theta)}{1 - F_2(\theta)} = 0.75. \end{cases}$$

If the resulting system has no solution, then go to Step 2.

Step 2. Assume now that both $q_1, q_3 < \theta$ and proceed like in Step 1. The two equations that result from matching the quartiles are now

$$\begin{cases} r\dfrac{F_1(q_1)}{F_1(\theta)} = 0.25, \\ r\dfrac{F_1(q_3)}{F_1(\theta)} = 0.75. \end{cases}$$

If, again, there is no corresponding solution, go to Step 3.

Step 3. Finally, assume that both $q_1, q_3 > \theta$ and proceed like in Step 1. When matching the quartiles, the two equations are

$$\begin{cases} r + (1-r)\dfrac{F_2(q_1) - F_2(\theta)}{1 - F_2(\theta)} = 0.25, \\ r + (1-r)\dfrac{F_2(q_3) - F_2(\theta)}{1 - F_2(\theta)} = 0.75. \end{cases}$$

If this system also has no solution, then try another distribution.

Remark 2.2. Once we have a solution from this second algorithm, we can use the first algorithm to improve it since this time we have some apriori information on θ.

2.2. SOME PROPERTIES OF A COMPOSITE TYPE II PARETO MODEL

As already mentioned in introduction, the choice of the Pareto distribution for f_2 is due to the fact of being a classical heavy-tailed distribution. In this section, we let f_1 be a general density function that we will particularize in Sections 3 and 4. In case we need this composite Pareto distribution to model e.g. insurance claims as discussed in Introduction, then we must choose f_1 to be less-heavy tailed than Pareto and defined only for positive values. Therefore, in Sections 3 and 4 we will illustrate the following results for f_1 being the Gamma and Weibull densities.

The composite Type II Pareto model will be developed in terms of the composite model (1.1), in which we will make use of a version of the generalized Pareto distribution (GPD) above the threshold value θ, i.e., we consider $f_2(x) = \frac{\alpha(\alpha\beta)^\alpha}{(\alpha\beta-\theta+x)^{\alpha+1}}$, where $\theta, \alpha, \beta > 0$ and $x > \theta$. This is also known as the Lomax or Type II Pareto distribution, see e.g., [2]. Now, letting $\gamma = \alpha\beta - \theta$, $\gamma > -\theta$, this density function may be written as

$$f_2(x) = \frac{\alpha\,(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}, \quad x > \theta.$$

Proposition 2.2. *The composite Type II Pareto density function is*

$$f(x) = \begin{cases} r\dfrac{f_1(x)}{F_1(\theta)}, & -\infty < x \le \theta, \\ (1-r)\dfrac{\alpha\,(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}, & \theta < x < \infty, \end{cases} \tag{2.6}$$

where the parameters $\alpha, \theta > 0$, $\gamma > -\theta$ and $0 \le r \le 1$ satisfy the continuity and differentiability conditions

$$\alpha + 1 = -\frac{(\gamma+\theta)\,f_1'(\theta)}{f_1(\theta)}, \tag{2.7}$$

$$r = \frac{\alpha f_1'(\theta) F_1(\theta)}{\alpha f_1'(\theta) F_1(\theta) - (\alpha+1)\,f_1^2(\theta)}. \tag{2.8}$$

Proof. The density function (2.6) results immediately from (1.1) and (2.1) by taking $f_2(x) = \frac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}$, $x > \theta$, and noticing that $F_2(\theta) = 0$. Hence, $f_1^*(x) = \frac{f_1(x)}{F_1(\theta)}$ and $f_2^*(x) = \frac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}$.

This density function has at least four unknown parameters: α, θ, γ and the mixing weight r. We now impose the continuity condition at the threshold point θ, i.e., $f(\theta - 0) = f(\theta + 0)$. This yields

$$r = \frac{\alpha F_1(\theta)}{\alpha F_1(\theta) + (\gamma + \theta) f_1(\theta)}. \tag{2.9}$$

We may ensure that the resulting density function is smooth if we also impose a differentiability condition at θ such that$f'(\theta - 0) = f'(\theta + 0)$. This restriction yields

$$r = \frac{\alpha(\alpha+1) F_1(\theta)}{\alpha(\alpha+1) F_1(\theta) - (\gamma+\theta)^2 f_1'(\theta)}. \tag{2.10}$$

From (2.9) and (2.10) we easily obtain the first condition, (2.7). Inserting $\gamma+\theta$ from this into (2.9) yields the expression (2.8) of r. □

Remark 2.3. Note that, because of conditions (2.7)–(2.8), the number of unknown parameters was reduced with two. We try to reduce it even more by imposing a second derivative requirement $f''(\theta - 0) = f''(\theta + 0)$, which leads to the condition

$$r\frac{f_1''(\theta)}{F_1(\theta)} = (1-r)\alpha(\alpha+1)(\alpha+2)\frac{1}{(\gamma+\theta)^3}.$$

Using now Proposition 2.1, we obtain the following

Corollary 2.1. *The cdf of the Type II Pareto model is*

$$F(x) = \begin{cases} r\dfrac{F_1(x)}{F_1(\theta)}, & -\infty < x \le \theta, \\ 1 - (1-r)\left(\dfrac{\gamma+\theta}{\gamma+x}\right)^\alpha, & \theta < x < \infty. \end{cases} \tag{2.11}$$

By Proposition 2.1 and $f_2^*(x) = \frac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}$, some calculation yields the following result.

Corollary 2.2. *The n-th order initial moment of the composite Type II Pareto distribution is*

$$E_n(f) = rE_n(f_1^*) + (1-r)\alpha\sum_{k=0}^{n}\binom{n}{k}\frac{(-\gamma)^{n-k}(\gamma+\theta)^k}{\alpha-k}, \quad \alpha > n. \tag{2.12}$$

3. COMPOSITE GAMMA-PARETO MODELS

3.1. THE COMPOSITE GAMMA-TYPE II PARETO MODEL

The composite Gamma-Type II Pareto model will be developed in terms of the mixture model (2.6). In the following, we denote by $\Gamma(\nu)=\int_0^\infty x^{\nu-1}e^{-x}\mathrm{d}x$ the Gamma function and by $\Gamma(\nu;t)=\int_0^t x^{\nu-1}e^{-x}\mathrm{d}x,\ \nu,t>0$, the lower incomplete Gamma function. For details on the Gamma distribution see e.g., [2].

PROPOSITION 3.1. *The composite Gamma-Type II Pareto density function is*

$$f(x)=\begin{cases} r\dfrac{\beta^\delta}{\Gamma(\delta;\beta\theta)}x^{\delta-1}e^{-\beta x}, & 0<x\le\theta,\\ (1-r)\dfrac{\alpha(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}, & \theta<x<\infty, \end{cases} \tag{3.1}$$

where the parameters $\beta,\delta,\alpha,\theta>0$, $\gamma>-\theta$ and $0\le r\le 1$ satisfy the continuity and differentiability conditions

$$\alpha+1=\frac{(\gamma+\theta)(\beta\theta-\delta+1)}{\theta}, \tag{3.2}$$

$$r=\frac{\alpha(\beta\theta-\delta+1)\Gamma(\delta;\beta\theta)}{\alpha(\beta\theta-\delta+1)\Gamma(\delta;\beta\theta)+(\alpha+1)(\beta\theta)^\delta e^{-\beta\theta}}. \tag{3.3}$$

Proof. The density function (3.1) immediately results from (2.6) by taking $f_1(x)=\frac{\beta^\delta}{\Gamma(\delta)}x^{\delta-1}e^{-\beta x}$, $x>0$ (i.e., a Gamma density function) and noting that $F_1(\theta)=\frac{\Gamma(\delta;\beta\theta)}{\Gamma(\delta)}$. Hence, the truncated Gamma density function is

$$f_1^*(x)=\frac{\beta^\delta}{\Gamma(\delta;\beta\theta)}x^{\delta-1}e^{-\beta x},\quad 0<x\le\theta.$$

Inserting now $f_1(\theta), f_1'(\theta)$ and $F_1(\theta)$ into (2.7) and (2.8), we easily obtain the first condition (3.2) and the expression (3.3) of r. □

Remark 3.1. Because of conditions (3.2)–(3.3), the number of unknown parameters was reduced from six to four (e.g., by expressing r and α in terms of β,γ,δ and θ). By Remark 2.3, the second derivative condition leads to

$$r\frac{\beta^\delta\theta^{\delta-3}}{\Gamma(\delta;\beta\theta)}\left((\delta-1)(\delta-2)-2\beta(\delta-1)\theta+\beta^2\theta^2\right)e^{-\beta\theta}=$$
$$=(1-r)\alpha(\alpha+1)(\alpha+2)\frac{1}{(\gamma+\theta)^3},$$

and using (3.2)–(3.3), this yields $\gamma(1-\delta) = \beta\theta^2$. Thus, the number of unknown parameters can be reduced at three.

Since $F_1(x) = \Gamma(\delta;\beta x)/\Gamma(\delta)$, applying Corollary 2.1 we immediately obtain:

COROLLARY 3.1. *The cdf of the composite Gamma-Type II Pareto model is*

$$F(x) = \begin{cases} r\dfrac{\Gamma(\delta;\beta x)}{\Gamma(\delta;\beta\theta)}, & 0 < x \le \theta, \\ 1-(1-r)\left(\dfrac{\gamma+\theta}{\gamma+x}\right)^{\alpha}, & \theta < x < \infty. \end{cases} \tag{3.4}$$

The next corollary results from Proposition 3.1 and Corollary 2.2.

COROLLARY 3.2. *The n-th order initial moment of the composite Gamma-Type II Pareto model is*

$$E_n(f) = r\frac{\Gamma(n+\delta;\beta\theta)}{\beta^n\,\Gamma(\delta;\beta\theta)} + (1-r)\,\alpha\sum_{k=0}^{n}\binom{n}{k}\frac{(-\gamma)^{n-k}(\gamma+\theta)^k}{\alpha-k},\ \alpha > n. \tag{3.5}$$

Remark 3.2. In the particular case when $\delta = n$ is a positive integer, the values $\Gamma(n;\cdot)$ involved in formula (3.5) can be evaluated recursively as

$$\Gamma(n+1;x) = n\Gamma(n;x) - x^n e^{-x}, \quad n \ge 1,\ x > 0, \tag{3.6}$$

with starting value $\Gamma(1;x) = 1-e^{-x},\ x > 0$. Applying this recursion successively also yields

$$\Gamma(n+1;x) = n! - \sum_{k=0}^{n}\frac{n!}{(n-k)!}x^{n-k}e^{-x}, \quad n \ge 1,\ x > 0,$$

but using directly recursion (3.6) seems to be a more efficient method in this case.

3.2. THE COMPOSITE GAMMA-PARETO MODEL

The composite Gamma-Pareto model results as a particular case of the composite Gamma-Type II Pareto model for $\gamma = 0$. From Proposition 3.1, Corollary 3.1 and Corollary 3.2, we easily obtain the following corollaries.

COROLLARY 3.3. *The composite Gamma-Pareto density is*

$$f(x) = \begin{cases} r\dfrac{\beta^{\delta}}{\Gamma(\delta;\beta\theta)}x^{\delta-1}e^{-\beta x}, & 0 < x \le \theta, \\ (1-r)\dfrac{\alpha\theta^{\alpha}}{x^{\alpha+1}}, & \theta < x < \infty, \end{cases} \tag{3.7}$$

where the parameters $\alpha, \beta, \delta, \theta > 0$ *and* $0 \leq r \leq 1$ *satisfy the continuity and differentiability conditions*

$$\alpha = \beta\theta - \delta, \tag{3.8}$$

$$r = \frac{(\beta\theta - \delta)\,\Gamma(\delta; \beta\theta)}{(\beta\theta - \delta)\,\Gamma(\delta; \beta\theta) + (\beta\theta)^{\delta} e^{-\beta\theta}}. \tag{3.9}$$

Remark 3.3. Note that because of the conditions (3.8)–(3.9), the number of unknown parameters can be reduced from five to three (e.g. we can express r and α in terms of β, δ and θ). When trying to reduce this number even more (based on the equality of the second derivatives), Remark 3.1 leads for $\gamma = 0$ to the condition $\beta\theta^2 = 0$, which is impossible.

COROLLARY 3.4. *The cdf of the composite Gamma-Pareto model is given by*

$$F(x) = \begin{cases} r\dfrac{\Gamma(\delta; \beta x)}{\Gamma(\delta; \beta\theta)}, & 0 < x \leq \theta, \\ 1 - (1 - r)\left(\dfrac{\theta}{x}\right)^{\beta\theta - \delta}, & \theta < x < \infty, \end{cases}$$

while its n-th order initial moment is

$$E_n(f) = r\frac{\Gamma(n + \delta; \beta\theta)}{\beta^n \Gamma(\delta; \beta\theta)} + (1 - r)\frac{\alpha\theta^n}{\alpha - n}, \quad \alpha > n. \tag{3.10}$$

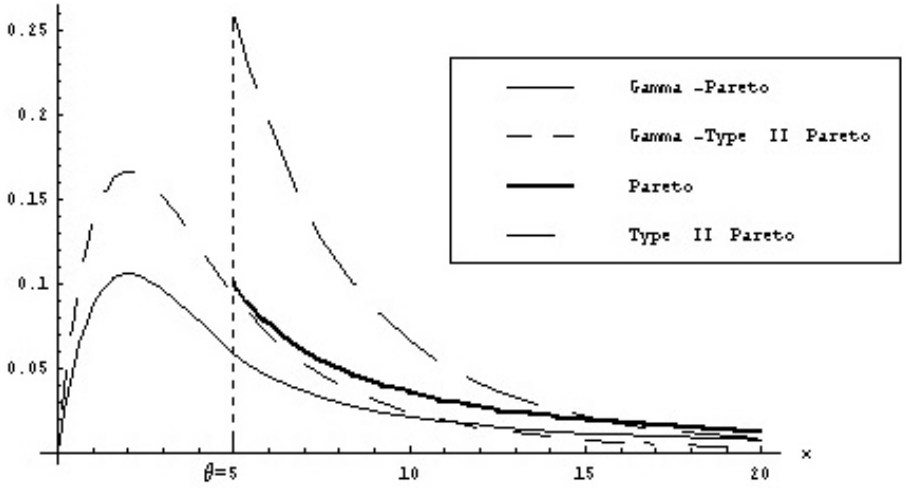

Fig. 1. Gamma-Pareto, Gamma-Type II Pareto, Pareto and Type II Pareto density curves for $\theta = 5$, $\beta = 0.5$, $\delta = 2$ and $\gamma = 20$.

In Figure 1 we display the density curves for the Type II Pareto and Pareto distributions, compared with the corresponding Gamma-Type II Pareto and Gamma-Pareto ones. Note that for this choice of parameters, the composite distributions are less heavy tailed.

Generating random values from the composite Gamma-Pareto distribution. In this case, both cdf-s F_1 and F_2 can be inversed. Then the inversion method presented in Subsection 2.1.2 becomes: if u is a value generated from the uniform distribution U(0,1), we obtain a value x generated from (3.7) as

If $u \leq r$ then, x is the solution of the equation $\Gamma(\delta; \beta x) = \frac{u}{r}\Gamma(\delta; \beta\theta)$.

If $u > r$ then, $x = \theta\left(\frac{1-r}{1-u}\right)^{1/\alpha}$.

It is not obvious how to inverse $\Gamma(\delta; \beta x)$, but this can be done with an appropriate software. For example, we used the function *InverseGammaReg* from Mathematica software (see, e.g., [8]).

Statistical inference. We start with the second algorithm presented in Subsection 2.1.3. From Remark 3.3, the number of unknown parameters is three, let them be β, δ and θ. Therefore, apart from the two quartiles q_1 and q_3, we need one more equation from the method of moments. We choose $\bar{x} = E_1(f)$, where $\bar{x}$ is the empirical mean of the data sample. By Corollary 3.4, the theoretical expected value of the density (3.7) is

$$E_1(f) = \frac{r\Gamma(\delta+1; \beta\theta)}{\beta\Gamma(\delta; \beta\theta)} + (1-r)\frac{\alpha\theta}{\alpha-1}, \quad \alpha > 1.$$

For simplicity, we denote $\xi = \beta\theta$ and replace $\theta = \frac{\xi}{\beta}$. Using also the expressions of r and α from (3.8) and (3.9), some calculation yield

$$E_1(f) = \frac{\xi-\delta}{\beta\left((\xi-\delta)\Gamma(\delta;\xi) + \xi^\delta e^{-\xi}\right)}\left(\Gamma(\delta+1;\xi) + \frac{\xi^{\delta+1}e^{-\xi}}{\xi-\delta-1}\right), \quad \alpha > 1.$$

Considering now, e.g., Step 1 from the second algorithm, where we assume that $q_1 < \theta < q_3$, we add two more equations for these quartiles. By Corollary 3.4, using also (3.8) and (3.9), these equations are

$$\text{(3.11)} \qquad \begin{cases} 0.25 = \dfrac{(\xi-\delta)\Gamma(\delta;\beta q_1)}{(\xi-\delta)\Gamma(\delta;\xi) + \xi^\delta e^{-\xi}}, \\ 0.25 = \dfrac{\xi^\xi e^{-\xi}}{\left((\xi-\delta)\Gamma(\delta;\xi) + \xi^\delta e^{-\xi}\right)(\beta q_3)^{\xi-\delta}}. \end{cases}$$

These two equations together with $\bar{x} = E_1(f)$ must be numerically solved.

The other two cases from Steps 2 and 3 result in a similar way.

When applying the first algorithm based on the ML method, the things are more complicated because one equation will involve the derivative of the

incomplete Gamma function, $\frac{\partial\Gamma(\delta;\xi)}{\partial\delta}$, which is not implemented in the usual software. More precisely, assuming that the data sample is ordered (as in Subsection 2.1.3) and that $x_m \leq \theta \leq x_{m+1}$ (information that we can get from e.g. the algorithm above), denoting $S_{ab} = \sum_{i=a}^{b} x_i$, $P_{ab} = \prod_{j=a}^{b} x_j$, after some calculation system (2.5) yields in our case

$$\begin{cases} \beta = \dfrac{\delta n + \xi\,(m-n)}{S_{1m}}, \\ \dfrac{n\xi^{\delta}e^{-\xi}}{(\xi-\delta)\left((\xi-\delta)\,\Gamma\,(\delta;\xi)+\xi^{\delta}e^{-\xi}\right)} + (n-m)\ln\dfrac{\xi}{\beta} - \ln P_{m+1n} = 0, \\ \dfrac{\xi^{\delta}e^{-\xi}\left(1+(\xi-\delta)\ln\xi\right)+(\xi-\delta)^2\,\frac{\partial\Gamma(\delta;\xi)}{\partial\delta}}{(\xi-\delta)\left((\xi-\delta)\,\Gamma\,(\delta;\xi)+\xi^{\delta}e^{-\xi}\right)} - \ln\beta - \dfrac{1}{n}\ln P_{1n} = 0. \end{cases}$$

We suggest to avoid the last equation that might create problems even in mathematical software, and combine the first two equations with, e.g., the equation $\bar{x} = E_1(f)$ from the method of moments.

Numerical example. We will now illustrate both estimation methods described above using $n = 5000$ generated data (by inversion method) form density (3.7). The real values of the parameters are

$$\theta = 3,\ \beta = 2,\ \delta = 4 \Rightarrow \alpha = 2,\ r = 0.760228,\ \xi = 6$$

while the empirical mean and quartiles needed for estimation are

$$\bar{x} = 2.678495, \quad q_1 = 1.326016, \quad q_3 = 2.879259.$$

We solved system (3.11) together with the equation $\bar{x} = E_1(f)$ using Mathematica software, and the solution obtained is

$$\tilde{\xi} = 6.196193, \quad \tilde{\beta} = 2.147827, \quad \tilde{\delta} = 4.201564,$$

hence $\tilde{\theta} = 2.884865$, $\tilde{\alpha} = 1.994629$, $\tilde{r} = 0.750968$.

Starting with this value of $\tilde{\theta}$, we applied the ML based algorithm suggested above to see if the parameters values can be improved. This was indeed the case, and we obtained the new slightly improved solutions

$$\hat{\xi} = 6.068690, \quad \hat{\beta} = 2.060728, \quad \hat{\delta} = 4.066182,$$

hence $\hat{\theta} = 2.944926$, $\hat{\alpha} = 2.002508$, $\hat{r} = 0.757894$. We mention that we started with $m = 3753$, but the final solution was obtained for $m = 3834$. Also, we used again Mathematica software to solve the resulting system.

We also applied the χ^2 test to check the distribution fitting for both sets of parameters, and to see which fit is better. The results are given in Table 1; to evaluate the cdf, we needed the Gamma function from Mathematica software. The corresponding 0.05 χ^2 quantile is 14.06712726, hence both

composite Gamma-Pareto distributions fit the generated data, but since the χ^2 distance is smaller for the ML parameters, their fit is better, as expected.

Table 1. Grouped data and χ^2 test (columns 2 results form the generated data; columns 3 and 5 are calculated using the Gamma-Pareto cdf for the two sets of estimated parameters)

Classes	Empirical freq., f_i	Theoretical freq., $\tilde{p}_i$ for $\tilde{\xi}, \tilde{\beta}, \tilde{\delta}$	$\frac{n\,(f_i-\tilde{p}_i)^2}{\tilde{p}_i}$	Theoretical freq., $\hat{p}_i$ for $\hat{\xi}, \hat{\beta}, \hat{\delta}$	$\frac{n\,(f_i-\tilde{p}_i)^2}{\tilde{p}_i}$
0.08–1	0.1288	0.127419	0.150×10^{-4}	0.129815	0.079×10^{-4}
1–1.3	0.1118	0.112023	0.004×10^{-4}	0.110839	0.083×10^{-4}
1.3–1.5	0.0818	0.082345	0.036×10^{-4}	0.081069	0.066×10^{-4}
1.5–1.8	0.1114	0.121984	9.183×10^{-4}	0.120045	6.226×10^{-4}
1.8–2	0.0838	0.074908	10.55×10^{-4}	0.073905	13.24×10^{-4}
2–2.5	0.1500	0.151721	0.195×10^{-4}	0.150808	0.043×10^{-4}
2.5–3	0.0990	0.099252	0.006×10^{-4}	0.100212	0.147×10^{-4}
3–4	0.1044	0.100570	1.459×10^{-4}	0.102158	0.492×10^{-4}
4–6	0.0692	0.071964	1.062×10^{-4}	0.072909	1.887×10^{-4}
6–50	0.0588	0.056956	0.597×10^{-4}	0.057387	0.348×10^{-4}
50–160	0.0010	0.000842	0.297×10^{-4}	0.000753	0.812×10^{-4}
$\sum$	1	χ^2 dist.=	11.77196	χ^2 dist. =	11.71531

4. COMPOSITE WEIBULL-PARETO MODELS

4.1. THE COMPOSITE WEIBULL-TYPE II PARETO MODEL

The composite Weibull-Type II Pareto results by taking f_1 to be a Weibull density function in (2.6). For details on the Weibull density see, e.g., [2].

PROPOSITION 4.1. *The composite Weibull-Type II Pareto density is*

$$f(x)=\begin{cases} r\dfrac{1}{1-e^{-(\theta/\tau)^\beta}}\cdot\dfrac{\beta}{\tau^\beta}x^{\beta-1}e^{-(x/\tau)^\beta}, & 0<x\le\theta,\\[2ex] (1-r)\dfrac{\alpha\,(\gamma+\theta)^\alpha}{(\gamma+x)^{\alpha+1}}, & \theta<x<\infty, \end{cases} \tag{4.1}$$

where the parameters $\beta,\tau,\alpha,\theta>0$, $\gamma>-\theta$ and $0\le r\le 1$ satisfy the continuity and differentiability conditions

$$\alpha+1=\frac{\gamma+\theta}{\theta}\big[\beta\,(\theta/\tau)^\beta-\beta+1\big], \tag{4.2}$$

$$r=\frac{\alpha\big[\beta\,(\theta/\tau)^\beta-\beta+1\big]\big[e^{(\theta/\tau)^\beta}-1\big]}{\alpha\big[\beta\,(\theta/\tau)^\beta-\beta+1\big]\big[e^{(\theta/\tau)^\beta}-1\big]+(\alpha+1)\,\beta\,(\theta/\tau)^\beta}. \tag{4.3}$$

Proof. The density function (4.1) immediately results from (2.6) by taking $f_1(x) = \frac{\beta}{\tau^\beta} x^{\beta-1} e^{-(x/\tau)^\beta}$, $\tau > 0$, $\beta > 0$, $x > 0$ (i.e., a Weibull density function) and using $F_1(\theta) = 1 - e^{-(\theta/\tau)^\beta}$. Hence,

$$f_1^*(x) = \frac{\beta}{\tau^\beta \left[1 - e^{-(\theta/\tau)^\beta}\right]} x^{\beta-1} e^{-(x/\tau)^\beta}, \quad x > 0.$$

Inserting $f_1(\theta), f_1'(\theta)$ and $F_1(\theta)$ into (2.7) and (2.8), we easily obtain the first condition (4.2) and then the expression (4.3) of r. □

Remark 4.1. Due to conditions (4.2)–(4.3), the number of unknown parameters was reduced from six to four (e.g., by expressing r and α in terms of β, τ, γ and θ). In the attempt to reduce this number even more, Remark 2.3 leads to the condition

$$r\beta (\theta/\tau)^\beta \frac{(\beta-1)(\beta-2) - 3\beta(\beta-1)(\theta/\tau)^\beta + \beta^2 (\theta/\tau)^{2\beta}}{\theta^3 \left[e^{(\theta/\tau)^\beta} - 1\right]} = \\ = (1-r)\alpha(\alpha+1)(\alpha+2)\frac{1}{(\gamma+\theta)^3},$$

and, using (4.2)–(4.3), $\gamma = \frac{\beta^2 \theta^{\beta+1}}{(1-\beta)\left(\tau^\beta + \beta\theta^\beta\right)}$.

Applying Corollary 2.1 with $F_1(x) = 1 - e^{-(x/\tau)^\beta}$ yields the following result.

COROLLARY 4.1. *The cdf of the composite Weibull-Type II Pareto model is*

$$F(x) = \begin{cases} r\dfrac{1 - e^{-(x/\tau)^\beta}}{1 - e^{-(\theta/\tau)^\beta}}, & 0 < x \le \theta, \\ 1 - (1-r)\left(\dfrac{\gamma+\theta}{\gamma+x}\right)^\alpha, & \theta < x < \infty. \end{cases} \tag{4.4}$$

From Corollary 2.2 and (4.1) we obtain next result.

COROLLARY 4.2. *The n-th order initial moment of the composite Weibull-Type II Pareto model is given by*

$$E_n(f) = r\frac{\tau^n}{1 - e^{-(\theta/\tau)^\beta}} \Gamma\left(\frac{n}{\beta} + 1; (\theta/\tau)^\beta\right) + \\ + (1-r)\alpha \sum_{k=0}^{n} \binom{n}{k} \frac{(-\gamma)^{n-k}(\gamma+\theta)^k}{\alpha - k}, \quad \alpha > n.$$

4.2. THE COMPOSITE WEIBULL-PARETO MODEL

The composite Weibull-Pareto model is obtained as the particular case of the composite Weibull-Type II Pareto model when $\gamma = 0$. The following corollaries easily result from the corresponding results in Subsection 4.1.

COROLLARY 4.3. *The composite Weibull-Pareto density is*

$$f(x) = \begin{cases} r\dfrac{1}{1-e^{-(\theta/\tau)^\beta}} \cdot \dfrac{\beta}{\tau^\beta} x^{\beta-1} e^{-(x/\tau)^\beta}, & 0 < x \le \theta, \\ (1-r)\dfrac{\alpha\theta^\alpha}{x^{\alpha+1}}, & \theta < x < \infty, \end{cases} \tag{4.5}$$

where the parameters $\alpha, \beta, \tau, \theta > 0$ and $0 \le r \le 1$ satisfy the continuity and differentiability conditions

$$\alpha = \beta\,(\theta/\tau)^\beta - \beta, \tag{4.6}$$

$$r = \frac{\alpha\left[e^{(\theta/\tau)^\beta} - 1\right]}{\alpha\left[e^{(\theta/\tau)^\beta} - 1\right] + \beta\,(\theta/\tau)^\beta}. \tag{4.7}$$

Remark 4.2. Using conditions (4.6)–(4.7), we reduced the number of unknown parameters from five to three (e.g. we can express r and α in terms of β, τ and θ). If we try to reduce it to only two parameters, letting $\gamma = 0$ in Remark 4.1, yields $\beta^2\theta^{\beta+1} = 0$, i.e., $\beta = 0$, which is impossible, or $\theta = 0$ when the model reduces to the classical Pareto.

COROLLARY 4.4. *The cdf of the composite Weibull-Pareto distribution is*

$$F(x) = \begin{cases} r\dfrac{1-e^{-(x/\tau)^\beta}}{1-e^{-(\theta/\tau)^\beta}}, & 0 < x \le \theta, \\ 1 - (1-r)\,(\theta/x)^\alpha, & \theta < x < \infty, \end{cases}$$

and its n-th order initial moment by

$$E_n(f) = r\frac{\tau^n}{1-e^{-(\theta/\tau)^\beta}}\Gamma\left(\frac{n}{\beta}+1; (\theta/\tau)^\beta\right) + (1-r)\,\frac{\alpha\theta^n}{\alpha-n}, \quad \alpha > n. \tag{4.8}$$

In the following, we plotted several density curves to give an idea of how these composite models look like. Figure 2 shows the density shapes for the composite Weibull-Pareto, Weibull-Type II Pareto, Pareto and Type II Pareto distributions, for similar parameters. Again, the composite distributions are less heavy-tailed. In Figure 3, we display several composite Weibull-Pareto densities for different values of parameter β, because it is known that this parameter gives the light-tailed or heavy-tailed character of the Weibull distribution (this is also the case for the composite distribution, as results from Figure 3).

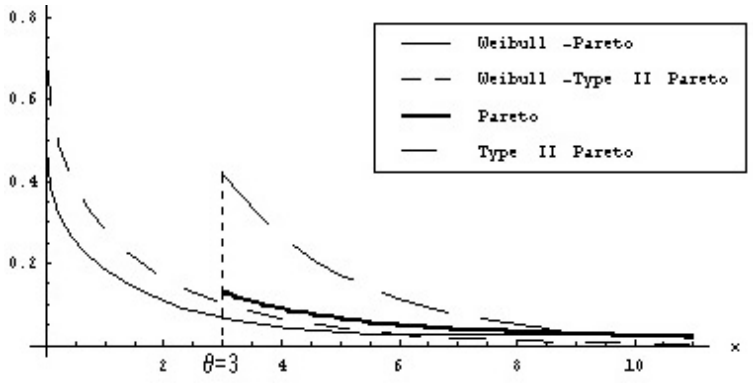

Fig. 2. Weibull-Pareto, Weibull-Type II Pareto, Pareto and Type II Pareto density curves for $\theta = 3$, $\beta = 0.9$, $\tau = 2$ and $\gamma = 20$.

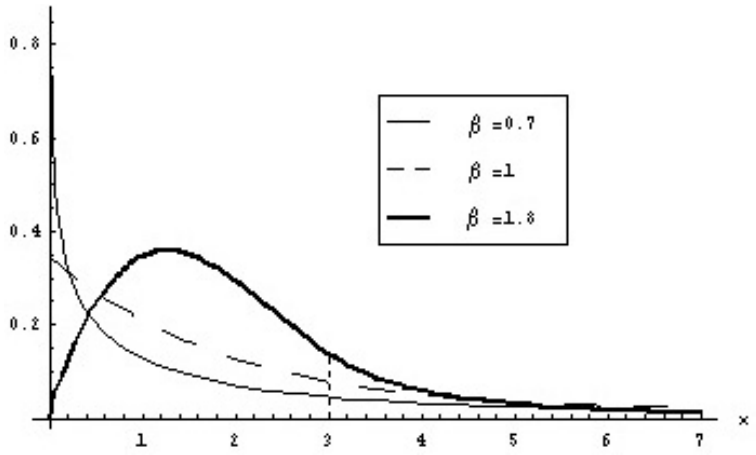

Fig. 3. Weibull-Pareto density curves for $\theta = 3$, $\tau = 2$ and different values of β.

Generating random values from the composite Weibull-Pareto distribution. Since, in this case, both cdf-s F_1 and F_2 can be inversed, the inversion method presented in Subsection 2.1.2 becomes: if u is a value generated from the uniform distribution U(0,1), we obtain a value x generated from (4.5) as

1. If $u \leq r$ then, $x = \tau \left[\ln \left(1 - \frac{u}{r} \left(1 - e^{-(\theta/\tau)^\beta} \right) \right)^{-1} \right]^{1/\beta}$.
2. If $u > r$ then, $x = \theta \left(\frac{1-r}{1-u} \right)^{1/\alpha}$.

Statistical inference. By Remark 4.2, the number of unknown parameters is three, i.e., β, τ and θ. As for the composite Gamma-Pareto model, we start again with the second algorithm given in Subsection 2.1.3. By Corollary 4.4, we have

$$E_1(f) = \frac{r\tau}{1-e^{-(\theta/\tau)^\beta}}\Gamma\left(\frac{1}{\beta}+1;(\theta/\tau)^\beta\right) + (1-r)\,\frac{\alpha\theta}{\alpha-1}, \quad \alpha>1,$$

and denoting $\zeta = (\theta/\tau)^\beta$, together with (4.6) and (4.7), we obtain

$$E_1(f) = \frac{\tau\,(\zeta-1)}{(\zeta-1)\,e^\zeta+1}\left(e^\zeta\Gamma\left(\frac{1}{\beta}+1;\zeta\right) + \frac{\beta\zeta^{1/\beta+1}}{\beta\,(\zeta-1)-1}\right), \quad \alpha>1.$$

Assuming now that $q_1<\theta<q_3$, we add two more equations for the quartiles. Using (4.6) and (4.7) in Corollary 4.4, these equations are

$$\text{(4.9)} \qquad \begin{cases} 0.25 = \dfrac{(\zeta-1)\,e^\zeta\left(1-e^{-(q_1/\tau)^\beta}\right)}{(\zeta-1)\,e^\zeta+1}, \\ 0.25 = \dfrac{\zeta^\zeta\tau^{\beta(\zeta-1)}}{((\zeta-1)\,e^\zeta+1)\,q_3^{\beta(\zeta-1)}}. \end{cases}$$

These two equations together with $\bar{x} = E_1(f)$ should be numerically solved. The other two cases (from Steps 2 and 3) can be similarly solved.

Using now the solution obtained from this algorithm as information for the first algorithm of Subsection 2.1.3, for an ordered data sample and $x_m \le \theta \le x_{m+1}$, denoting as before $S_{ab} = \sum_{i=a}^{b} x_i$, $P_{ab} = \prod_{j=a}^{b} x_j$, system (2.5) yields after some calculation

$$\begin{cases} \dfrac{1}{\tau^\beta}\sum_{i=1}^{m} x_i^\beta\,(\ln x_i - \ln\tau) - \dfrac{n}{\beta} - \ln P_{1m} + (\zeta-1)\ln P_{m+1n} - \\ \qquad\qquad - ((n-m)\,(\zeta-1)-m)\ln\tau = 0, \\ \dfrac{n\zeta}{(\zeta-1)\,((\zeta-1)\,e^\zeta+1)} + (n-m)\,(\ln\zeta+\beta\ln\tau) - \beta\ln P_{m+1n} = 0, \\ \dfrac{1}{\tau^\beta}\sum_{i=1}^{m} x_i^\beta + (n-m)\,(\zeta-1) - m = 0, \end{cases}$$

This system will be solved numerically.

Numerical example. To illustrate the estimation method presented before, we generated a data sample of volume $n = 5000$ from density (4.5), using the generating method based on inversion. The real values of the parameters are

$$\theta = 10,\ \beta = 2,\ \tau = 7 \Rightarrow \alpha = 2.0816327,\ r = 0.7735207,\ \zeta = 2.0408163,$$

while the empirical mean and quartiles needed for estimation are

$$\bar{x} = 8.3822291, \quad q_1 = 4.0635264, \quad q_3 = 9.6417444.$$

We solved the system (4.9) plus equation $\bar{x} = E_1(f)$ using Mathematica software, and the solution obtained is

$$\tilde{\zeta} = 2.1068866, \quad \tilde{\beta} = 1.9738183, \quad \tilde{\tau} = 7.1812994,$$

hence, $\tilde{\theta} = 10.4753873$, $\tilde{\alpha} = 2.1847930$, $\hat{r} = 0.7914281$. To solve the system, we had to indicate Mathematica two starting values for each unknown parameter; by varying these starting values, we noticed that Mathematica gave the same solution even if the starting values were quite far away from the real ones.

We also applied the χ^2 test to check the distribution fitting for these parameters values, and the results are given in Table 2. The χ^2 distance is 11.933639281, while the corresponding 0.05 χ^2 quantile is 14.06712726, hence the composite Weibull-Pareto distribution fits our data.

Table 2. Grouped data and χ^2 test (columns 2 results form the data sample; column 3 is calculated using the Weibull-Pareto cdf)

Classes	Empirical freq., f_i	Theoretical freq., p_i	$\frac{n(f_i - p_i)^2}{p_i}$
0.1-1.5	0.0376	0.039843044	0.631383258
1.5-3	0.1104	0.107292352	0.450054395
3-4	0.0942	0.096158030	0.199353073
4-5	0.1162	0.105197008	5.754243350
5-6	0.1096	0.105500528	0.796473176
6-7	0.0926	0.098635108	1.846326578
7-8	0.0822	0.086794872	1.216249680
8-10	0.1258	0.129602882	0.557931717
10-14	0.1216	0.120102166	0.093399949
14-30	0.0878	0.089741748	0.210068558
30-285	0.0220	0.021132264	0.178155548
$\sum$	1	χ^2 dist. =	11.933639281

5. CONCLUSIONS

In this paper, we generalize the composite models studied in connection with insurance data by Cooray and Ananda [1] and Scollnik [6]. We think that the general models studied in Section 2 can be used on any statistical data that present a shape corresponding to a combination of two densities in the form (1) or, equivalently, (2). If, starting from a certain threshold on, the data display a heavy-tailed trend, then the particular composite Pareto distributions presented in Sections 3 and 4 could be adequate fits.

REFERENCES

[1] K. Cooray and M.A. Ananda, *Modeling actuarial data with a composite Lognormal-Pareto model*. Scand. Actuar. J. **5** (2005), 321–334.
[2] N.L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions*. Wiley, New York, 1994.
[3] R. Kaas, M. Goovaerts, M. Denuit and J. Dhaene, *Modern Actuarial Risk Theory*. Kluwer, Boston, 2001.
[4] S.A. Klugman, H.H. Panjer and G.E. Willmot, *Loss Models: from Data to Decisions* (2nd Edition). Wiley, New York, 2004.
[5] V. Preda and R. Ciumara, *On composite models: Weibull-Pareto and Lognormal-Pareto. A comparative study*. Rom. J. Econ. Forecast. **3** (2006), 32–46.
[6] D.P.M. Scollnik, *On composite Lognormal-Pareto models*. Scand. Actuar. J. **1** (2007), 20–33.
[7] S. Teodorescu and R. Vernic, *Some composite Exponential-Pareto models for actuarial prediction*. Rom. J. Econ. Forecast. **12** (2009), 82–100.
[8] Mathematica Web-site: `www.wolfram.com`.

Received 1 January 2011

"Nicolae Titulescu" University
Faculty of Economic Sciences
185 Calea Văcăreşti
040051 Bucharest, Romania
cezarina_teodorescu@yahoo.com

"Ovidius" University of Constanţa
Faculty of Mathematics and Computer Science
124 Mamaia Blvd.
900527 Constanţa, Romania

and

Institute for Mathematical Statistics
and Applied Mathematics
Calea 13 Septembrie 13
050711 Bucharest, Romania
rvernic@univ-ovidius.ro