MODELLING INSURANCE DATA WITH THE PARETO ARCTAN DISTRIBUTION

BY

EMILIO GÓMEZ–DÉNIZ AND ENRIQUE CALDERÍN–OJEDA

ABSTRACT

In this paper, a new methodology based on the use of the inverse of the circular tangent function that allows us to add a scale parameter (say α) to an initial survival function is presented. The latter survival function is determined as limiting case when α tends to zero. By choosing as parent the classical Pareto survival function, the Pareto ArcTan (PAT) distribution is obtained. After providing a comprehensive analysis of its statistical properties, theoretical results with reference to insurance are illustrated. Its performance is compared, by means of the well-known Norwegian fire insurance data, with other existing heavy-tailed distributions in the literature such as Pareto, Stoppa, Shifted Lognormal, Inverse Gamma and Fréchet distributions.

KEYWORDS

Limited expected value, loss, Pareto distribution, shifted lognormal distribution, Stoppa sistribution, threshold.

1. INTRODUCTION

Different methods for generalizing discrete or continuous probability density functions have been proposed in the statistical literature. Most of these procedures are based on the idea of incorporating a new parameter to a classical distribution, as occurs with the class of max-stable distributions (see Sarabia and Castillo (2005) for a revision of this class of distributions) and the Marshall and Olkin family of distributions proposed in Marshall and Olkin (1997). The probabilistic models derived from these and other methodologies generally exhibit flexibility and a collection of different shapes and hazard functions for modelling data sets of diverse nature. Additionally, they encompass classical distributions for particular values of the new parameter attached to the initial family.

Astin Bulletin 45(3), 639-660. doi: 10.1017/asb.2015.9

In this work, a new method to add a parameter to a family of distributions is proposed after making a change of variable in the truncated Cauchy distribution. As a consequence of this, a class of probabilistic models that includes an additional scale parameter $\alpha \neq 0$ and also the inverse of the circular tangent function ($\tan^{-1}(\cdot)$) in its analytical expression is obtained. To the best of our knowledge, the methodology introduced in this work is a novelty and it seems suitable for being applied to different probabilistic families. In this work, the Pareto distribution is considered.

In general insurance, only a few large claims arising in the portfolio represent the largest part of the payments made by the insurance company. Appropriate estimation of these extreme events is crucial for the practitioner to correctly assess insurance and reinsurance premiums. On this subject, the single parameter Pareto distribution (Arnold (1983), Brazauskas and Serfling (2003), Rytgaard (1990), among others) has been traditionally considered as a suitable claim size distribution in relation to rating problems. Concerning this, the single parameter Pareto distribution, apart from its nice properties, provides a good depiction of the random behavior of large losses (e.g. the right tail of the distribution). Particularly, when calculating deductibles and excess-of-loss levels for reinsurance, the simple Pareto distribution has been demonstrated convenient, see for instance Boyd (1988), Mata (2000) and Klugman *et al.* (2008), among others.

In this work, an extension of the Pareto distribution, by using the methodology described above, is derived. This new model provides a more accurate description of large losses in terms of high quantiles, in the context of the Norwegian fire insurance data set, than other distributions traditionally proposed in the actuarial literature such as the classical Pareto, shifted Lognormal, Burr, Inverse Gamma or Fréchet among others. Firstly, a comprehensive treatment of its mathematical properties is provided. In this sense, expressions for the moments, variance, cumulative distribution function, asymptotic ruin function, VaR, TVaR and limited expected values, among other properties are derived. Estimation of the parameter of this distribution can be easily calculated by the maximum likelihood method by using numerical search of the maximum. Although the moments method is also possible by means of numerical techniques, this is not shown in this work.

Finally, the performance of the model introduced in this manuscript is tested by means of the well-known Norwegian fire insurance portfolio data. The new distribution proposed in this work will be compared to other heavy-tailed distributions existing in the literature: Pareto, shifted Lognormal, Burr, Loggamma, Fréchet, Inverse Gamma and Stoppa (a well-known generalization of the Pareto distribution that belongs to the max-stable distribution (see Kleiber and Kotz, 2003)) distributions. As will be seen later, the new distribution introduced in this manuscript, denoted as Pareto ArcTan (PAT), outperforms all these probabilistic families in terms of different measures of model validation and it also provides a closer estimation of high quantiles.

The plan for the paper is as follows. In the next section, the new methodology to incorporate a scale parameter to a family of distributions is described in detail. Next, in Section 3, this approach is applied to the classical Pareto distribution to derive the Pareto ArcTan model. Then, some statistical properties of the model are given. In Section 4, several theoretical results relevant to insurance are presented. Afterward, in Section 5, numerical applications are illustrated based upon the Norwegian fire insurance set of data. Here, the new generalization of Pareto distribution is compared with the models mentioned above under different measures of goodness-of-fit. Also estimation of high quantiles is provided for some of the models considered. Besides, score equations of PAT distribution are shown. The final section contains discussion and conclusion.

2. Main result

The half-Cauchy distribution (Jacob and Jayakumar (2012)) truncated at $\alpha > 0$ has probability density function given by

$$f(y) = \frac{1}{\tan^{-1}\alpha}\frac{1}{1+y^2}, \quad 0 < y < \alpha. \tag{1}$$

In the latter expression, $\tan^{-1}$ is the inverse of the circular tangent function. Let us consider now the transformation $y = \alpha\bar{F}(x)$, where $\bar{F}(x)$ is the survival function of a random variable X with support in $[a, b]$ and where a and b can be finite or non-finite. Then, the corresponding probability density function of the random variable X obtained from (1) results

$$f(x;\alpha) = \frac{1}{\tan^{-1}\alpha}\frac{\alpha f(x)}{1+[\alpha\bar{F}(x)]^2}, \tag{2}$$

for $a \le x \le b$ and $\alpha > 0$. The survival function of X, which is obtained from (2) by integrating, is given by

$$\bar{F}(x;\alpha) = \frac{\tan^{-1}(\alpha\bar{F}(x))}{\tan^{-1}\alpha}. \tag{3}$$

Furthermore, it is simple to see that (2) and (3) are the proper probability density function and survival function, respectively, when the support of the parameter α is extended to $(-\infty, \infty)$ except for zero. In this case, we get that $\bar{F}(x;\alpha) = \bar{F}(x;-\alpha)$. Additionally, by taking in (3) limit when the parameter α tends to zero and applying L'Hospital's rule, it is straightforward to derive that the parent survival function, $\bar{F}(x)$, is obtained as a special case, i.e. $\bar{F}(x;\alpha) \to \bar{F}(x)$ when $\alpha \to 0$. Thus, the methodology proposed here can be considered as a mechanism for adding a scale parameter to a parent survival function and, therefore, a means of obtaining a more flexible survival function. In particular, the case where $\bar{F}(x)$ is replaced by the survival function of the classical Pareto distribution will be considered in this manuscript.

Alternatively, the expression $\bar{F}(x;\alpha)$ provided in (3) can be obtained by using the following representation of the $\tan^{-1}$ function (see Castellanos (1988)):

$$\tan^{-1} z = \frac{z}{1+z^2}\,{}_2F_1\left(1,1;\frac{3}{2},\frac{z^2}{1+z^2}\right) = \sum_{n=0}^{\infty}\frac{2^{2n}(n!)^2}{(2n+1)!}\frac{z^{2n+1}}{(1+z^2)^{n+1}}.$$

Here ${}_2F_1(a,b;c,z)$ represents the hypergeometric function which has the integral representation

$${}_2F_1(a,b;c,z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\,dt,$$

and where $\Gamma(\cdot)$ is the Euler gamma function.

The hazard rate function defined by $h(x) = f(x)/\bar{F}(x)$ is an important quantity to characterize life phenomena. Derivation of this function is straightforward from (2) and (3) and is given by

$$h(x;\alpha) = \frac{1}{\tan^{-1}[\alpha\bar{F}(x)]}\frac{\alpha f(x)}{1+[\alpha\bar{F}(x)]^2}. \tag{4}$$

Although in this paper it has only been considered the transformation of a survival function by using the $\tan^{-1}$ function, it is also possible to obtain a family of distributions by replacing this function by the tan function. Nevertheless, the new family of distributions obtained seems not very flexible. In fact, under this alternative transformation, the new Pareto distribution attains its modal value close to the lower bound of its support. Additionally, the parameter α has a more restrictive domain if we wish to ensure that the transformation provides a genuine survival function.

3. The Pareto case

3.1. Genesis and properties

In this section, a new generalization of the Pareto distribution is obtained by using the methodology proposed above. In this regard, observe that when (3) is applied to the classical Pareto distribution with survival function $\bar{F}(x) = (\frac{\sigma}{x})^{\theta}$, $x \geq \sigma$, $\sigma > 0$, the new survival function is attained and it is provided by

$$\bar{F}(x;\alpha) = \frac{\tan^{-1}(\alpha(\sigma/x)^{\theta})}{\tan^{-1}\alpha}, \quad x \geq \sigma. \tag{5}$$

Then, the corresponding probability density function is

$$f(x;\alpha) = \frac{1}{\tan^{-1}\alpha}\frac{\alpha\theta\sigma^{\theta}x^{\theta-1}}{(\alpha\sigma^{\theta})^2 + x^{2\theta}}, \quad x \geq \sigma. \tag{6}$$

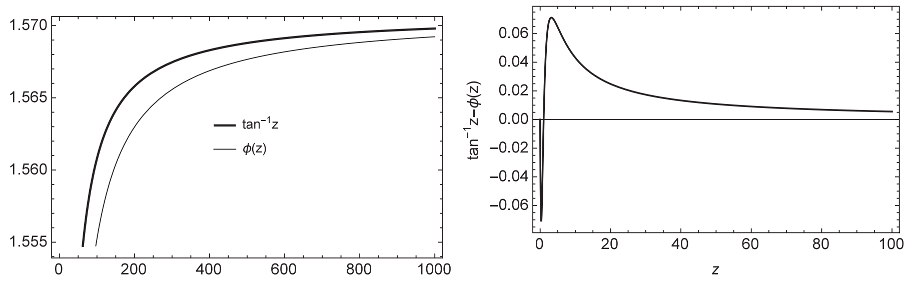

FIGURE 1: Graphs of $\tan^{-1}(z)$ and $\phi(z)$ functions (left) and error function (right).

Undoubtedly, the single parameter Pareto distribution is one of the most attractive distributions in statistics; a power law probability distribution found in a large number of real-world situations inside and outside the field of economics. Furthermore, it is usually used as a basis for excess of loss quotations as it gives a pretty good description of the random behavior of large losses. The probability density function (6) includes the Pareto distribution as a limiting case when $\alpha \to 0$. A tentative interpretation of the parameters of this distribution can be given in the context of insurance setting and, in particular in reinsurance context. In this field, there are many probability distributions which can be used for modelling single loss amounts. Then, if the loss is assumed to follow the probability density function (6), we have that σ determines the minimum loss amount and θ defines the tail behavior of the distribution.

Also note that approximation of the $\tan^{-1}$ function can be obtained by using second- and third-order polynomials and simple rational functions (see Rajan *et al.* (2006) for details). In this sense, it is verified that

$$\tan^{-1}\left(\frac{1+x}{1-x}\right) \approx \frac{\pi}{4}(x+1), \quad -1 \leq x \leq 1$$

and after appropriate change of variable we have that

$$\tan^{-1} z \approx \frac{z\pi}{2(1+z)}, \quad z \geq 0. \tag{7}$$

Figure 1 shows both functions, $\tan^{-1}(z)$ and $\phi(z) = z\pi/(2(1+z))$, for values $z > 0$ together with the error function calculated for their difference. As can be observed, the approximation is reasonably good reaching a maximum at $z = 3.1904$ and a minimum at $z = 0.3134$ with $|\tan^{-1} z - \phi(z)| \leq 0.0711$.

Applying (7) to (3), for $\alpha > 0$, we get after some algebra that

$$\bar{F}(x;\beta) \approx \frac{(1+\alpha)\bar{F}(x)}{1+\alpha\bar{F}(x)}, \quad \alpha > 0, \tag{8}$$

which is related to the family of distributions proposed by Marshall and Olkin (1997). Then, the mechanism introduced here is more general than the one

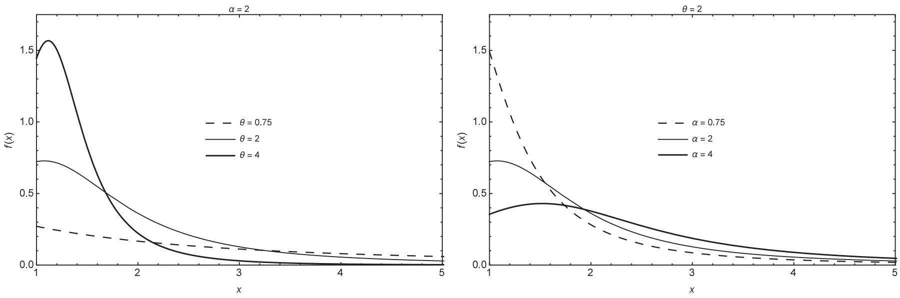

FIGURE 2: Probability density function of PAT distribution for selected values of parameters α and θ assuming $\sigma = 1$.

proposed by these authors. Furthermore, for this particular case, the family of distributions presented in (3) is geometric-minimum stable (see Marshall and Olkin (1997)). In this sense, a simple interpretation of the new model is given as follows. Let us suppose that $\{X_i\}_{i=1}^n$ are independent and identically distributed random variables with cumulative distribution function $F(x)$, where n is random and it follows the probability mass function

$$\Pr(N = n) = \frac{1}{1+\alpha}\left(\frac{\alpha}{1+\alpha}\right)^{n-1}, \quad n = 1, 2, \ldots \tag{9}$$

(i.e. the geometric distribution); then it is simple to show that the marginal survival function of $X = \min\{X_1, X_2, \ldots, X_n\}$ is given by (8). For example, we might think of a random sequence of losses where the time between claims has a mean of α^{-1}. Likewise the new model can be considered a way to describe the time between claims occurrence.

Henceforward, when a random variable X follows the probability density function given in (6) we will write $X \sim \text{PAT}(\alpha, \theta, \sigma)$. The shape of density (6) is shown in Figure 2 for a few values of the parameters, with α fixed and θ varying in the plot on the left and θ fixed and α changing in the graph on the right side. Note that only positive values of α have been considered since for negative values the density is the same.
The rth moment about zero is given in the next result.

Proposition 1. *The rth moment about the origin of the PAT distribution is given by*

$$E(X^r) = \frac{\alpha\theta\sigma^r}{(\theta - r)\tan^{-1}\alpha}\, {}_2F_1\left(1, \frac{\theta - r}{2\theta}; \frac{3\theta - r}{2\theta}; -\alpha^2\right), \quad \theta > r, \tag{10}$$

Proof. We have that

$$E(X^r) = \frac{\alpha\theta\sigma^\theta}{\tan^{-1}\alpha}\int_\sigma^\infty \frac{x^{r+\theta-1}}{(\alpha\sigma^\theta)^2 + x^{2\theta}}\, dx.$$

https://doi.org/10.1017/asb.2015.9 Published online by Cambridge University Press

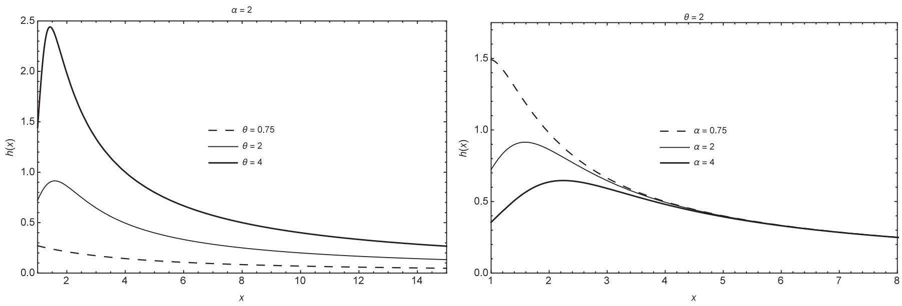

FIGURE 3: Failure rate function of PAT distribution for selected values of parameters α and θ assuming $\sigma = 1$.

Now, by making the change of variable $x = \sigma/z^{1/(2\theta)}$ we get the result after some algebra. ■

The fact that the moments in (10) can be expressed in terms of the hypergeometric function guarantees convergence and existence. Furthermore, the use of software, such as Matlab or Mathematica, that incorporates this special function facilitates its practical implementation.

In particular, the mean takes the form

$$\mu = E(X) = \frac{\alpha\theta\sigma}{(\theta-1)\tan^{-1}\alpha}\, {}_2F_1\left(1, \frac{\theta-1}{2\theta}, \frac{3\theta-1}{2\theta}, -\alpha^2\right), \quad \theta > 1. \quad (11)$$

From (5), the quantile function x_γ is simply derived

$$x_\gamma = \sigma\left[\frac{1}{\alpha}\tan(\gamma\tan^{-1}\alpha)\right]^{-1/\theta},$$

from which the median can be easily obtained.

Besides, the mode, which can be obtained by differentiating (6) with respect to the variable x, is given by

$$x_{Mo} = \left[\frac{(\theta-1)(\sigma^\theta\alpha)^2}{1+\theta}\right]^{1/2\theta}.$$

Finally, the hazard rate function for the PAT distribution, which is obtained from (2), (3) and (4), has been plotted for the same values of parameters as considered in the previous figure. This is shown in Figure 3.

As can be seen the hazard rate function is either monotonically decreasing or increasing and decreasing.

https://doi.org/10.1017/asb.2015.9 Published online by Cambridge University Press

3.2. Stochastic ordering

The quantity given in (5) represents the proportion of claims larger than a given x value; hence, examination of the statistical ordering with respect to the parameter α could be convenient to better explain the claim size distribution. The PAT distribution can be ordered by some stochastic orders according to the value of the parameter α (assumed positive) in terms of the likelihood ratio order, which is defined as follows.

Definition 1. *Let X_1 and X_2 be continuous random variables with densities f_1 and f_2, respectively, such that*

$$\frac{f_2(x)}{f_1(x)} \text{ is non-decreasing over the union of the supports of } X_1 \text{ and } X_2.$$

Then X_1 is said to be smaller than X_2 in the likelihood ratio order (denoted by $X_1 \leq_{LR} X_2$).

Theorem 1. *Let X_1 and X_2 be two PAT random variables with density functions $f(x|\alpha_1, \theta, \sigma)$ and $f(x|\alpha_2, \theta, \sigma)$, respectively. If $0 < \alpha_1 \leq \alpha_2$ then $X_1 \leq_{LR} X_2$.*

Proof. Note that the ratio

$$h(x) = \frac{f(x|\alpha_2, \theta, \sigma)}{f(x|\alpha_1, \theta, \sigma)} = \frac{\alpha_2 \tan^{-1} \alpha_1 \left(\alpha_1^2 \sigma^{2\theta} + x^{2\theta}\right)}{\alpha_1 \tan^{-1} \alpha_2 \left(\alpha_2^2 \sigma^{2\theta} + x^{2\theta}\right)}$$

is non-decreasing if and only if $h'(x) \geq 0$ for $x \in (\sigma, \infty)$. Some calculations show that

$$h'(x) = \frac{2\alpha_2 \theta \tan^{-1} \alpha_1 (\alpha_2 - \alpha_1)(\alpha_1 + \alpha_2) \sigma^{2\theta} x^{2\theta - 1}}{\alpha_1 \tan^{-1} \alpha_2 \left(\alpha_2^2 \sigma^{2\theta} + x^{2\theta}\right)^2}.$$

Now, taking into account that if $0 < \alpha_1 \leq \alpha_2$ then it is easy to see that for all $\theta \in \mathbb{R}^+$ and $x \geq \sigma$ imply $h'(x) \geq 0$ and the result holds. ■

4. Theoretical results in insurance

In the following, several theoretical results related to insurance for the PAT distribution are derived.

4.1. Heavy-tailed distributions

The use of heavy right-tailed distributions is of vital importance in general insurance. In this regard, Pareto and Lognormal distributions have been employed to model losses in motor third-party liability insurance, fire insurance or catastrophe insurance. It is already known that any probability distribution, that is specified through its cumulative distribution function $F(x)$ on the real line, is

https://doi.org/10.1017/asb.2015.9 Published online by Cambridge University Press

heavy right-tailed (see Rolski *et al.* (1999)) if $\limsup_{x\to\infty} -\log \bar{F}(x)/x = 0$. This is simply verified when $\bar{F}(x)$ is replaced by (5). Therefore, we have that $\lim_{x\to\infty} e^{sx}\bar{F}(x;\alpha) = \infty$ for $x > 0$.

Therefore, as a long-tailed distribution is also heavy right-tailed, the distribution introduced in this manuscript is also heavy right-tailed.

Another important issue in extreme value theory is the regular variation (see Bingham *et al.* (1987)). This concept is formalized in the following definition.

Definition 2. *A distribution function is called regular varying at infinity with index* $-\beta$ *if*

$$\lim_{x\to\infty} \frac{\bar{F}(tx)}{\bar{F}(x)} = t^{-\beta},$$

where the parameter $\beta \geq 0$ *is called the tail index.*

Next theorem establishes that the survival function given in (5) is a regular variation Lebesgue measure.

Theorem 2. *The survival function given in (5) is a survival function with regularly varying tails.*

Proof. Let us firstly consider the survival function given in (5). Then, after simple computations it is obtained that

$$\lim_{x\to\infty} \frac{\bar{F}(tx;\alpha)}{\bar{F}(x;\alpha)} = t^{-\theta},$$

and taking into account that $\theta \geq 0$ the statement of the theorem follows. ■

As a consequence of this result, we have that if $X, X_1, \ldots, X_n$ are iid random variables with the common survival function given by (5) and $S_n = \sum_{i=1}^{n} X_i,\ n \geq 1$, then

$$\Pr(S_n > x) \sim \Pr(X > x) \quad \text{as} \quad x \to \infty.$$

Therefore, if $P_n = \max_{i=1,\ldots,n} X_i,\ n \geq 1$, we have that

$$\Pr(S_n > x) \sim n \Pr(X > x) \sim \Pr(P_n > x).$$

This means that for large x the event $\{S_n > x\}$ is due to the event $\{P_n > x\}$. Therefore, exceedances of high thresholds by the sum S_n are due to the exceedance of this threshold by the largest value in the sample.

On the other hand, let the random variable X represent either a policy limit or reinsurance deductible (from an insurer's perspective); then the limited expected value function L of X with cdf $F(x)$ is defined by

$$L(x) = E[\min(X, x)] = \int_0^x y\, dF(y) + x\bar{F}(x),$$

which is the expectation of the cdf $F(x)$ truncated at this point. In other words, it represents the expected amount per claim retained by the insured on a policy with a fixed amount deductible of x.

For the PAT distribution proposed here this amount is given by

$$\begin{aligned}L(x) &= \frac{\theta(x/\sigma)^{\theta}}{x\alpha(\theta+1)\tan^{-1}\alpha}\, {}_2F_1\left(1, \frac{\theta+1}{2\theta}; \frac{3\theta+1}{2\theta}; -\frac{(x/\sigma)^{2\theta}}{\alpha^2}\right)\\ &+ \frac{x}{\tan^{-1}\alpha}\tan^{-1}\left(\alpha\left(\frac{\sigma}{x}\right)^{\theta}\right)\\ &- \frac{\sigma^{\theta+1}}{\alpha(\theta+1)\tan^{-1}(\alpha)}\, {}_2F_1\left(1, \frac{\theta+1}{2\theta}; \frac{3\theta+1}{2\theta}; -\frac{1}{\alpha^2}\right).\end{aligned}$$

The value at risk (VaR) is defined as the amount of capital required to ensure that the insurer does not become insolvent with a high degree of certainty. The VaR of a random variable X which follows the PAT distribution is the q quantile and it is given by

$$\mathrm{VaR}(X; q) = \sigma\left(\frac{\alpha}{\tan((1-q)\tan^{-1}\alpha))}\right)^{1/\theta}.$$

The use of the VaR is questionable due to the lack of subadditivity, for that reason the expected loss given that the loss exceeds the q quantile of the distribution of X, the tail value at risk (TVaR), is considered. Then, if X follows a PAT distribution, for any quantile q the tail value at risk is given by

$$\begin{aligned}\mathrm{TVaR}(X; q) &= \frac{1}{1-q}\int_q^1 \mathrm{VaR}(x; q)\, dq\\ &= -\frac{\sigma\theta\alpha^{1/\theta}\tan^{\frac{\theta-1}{\theta}}\left(\tan^{-1}\alpha - q\tan^{-1}\alpha\right)}{(\theta-1)(q-1)\tan^{-1}(\alpha)}\\ &\quad\times {}_2F_1\left(1, \frac{\theta-1}{2\theta}; \frac{3\theta-1}{2\theta}; -\tan^2\left(\tan^{-1}\alpha - q\tan^{-1}\alpha\right)\right).\end{aligned}$$

The probability density function given in (6) can also be applied in rating excess-of-loss reinsurance as it can be seen in the next result.

Proposition 2. *Let X be a random variable denoting the individual claim size taking values only for individual claims greater than d. Let us also assume that X follows the probability density function (6), and then the expected cost per claim to the reinsurance layer when the loss in excess of m subject to a maximum of l is*

https://doi.org/10.1017/asb.2015.9 Published online by Cambridge University Press

given by

$$E[\min(l, \max(0, X-m))] = \sigma^{-\theta}\left(\alpha(\theta+1)\tan^{-1}\alpha\right)^{-1}$$

$$\times\left[\theta(L+m)^{\theta+1}\,{}_2F_1\left(1, \frac{\theta+1}{2\theta}; \frac{3\theta+1}{2\theta}; -\frac{\sigma^{-2\theta}(L+m)^{2\theta}}{\alpha^2}\right)\right.$$

$$-\ \theta m^{\theta+1}\,{}_2F_1\left(1, \frac{\theta+1}{2\theta}; \frac{3\theta+1}{2\theta}; -\frac{\sigma^{-2\theta}m^{2\theta}}{\alpha^2}\right)$$

$$+\ \alpha(\theta+1)\sigma^{\theta}\left(-m\cot^{-1}\left(\alpha\sigma^{\theta}(L+m)^{-\theta}\right)+m\cot^{-1}\left(\alpha\sigma^{\theta}m^{-\theta}\right)\right.$$

$$\left.\left.+\ L\tan^{-1}\left(\alpha\left(\frac{\sigma}{L+m}\right)^{\theta}\right)\right)\right].$$

Proof. The result follows by taking into account that

$$E[\min(l, \max(0, X-m))] = \int_m^{m+l}(x-m)f(x)\,dx + l\bar{F}(m+l),$$

from which we get the result after some algebra. ■

4.2. Approximating the ruin function

The integrated tail distribution (also known as equilibrium distribution) is an important distribution that often appears in insurance and many other applied probability models. Let $\bar{F}(x)$ be the survival function given in (5), and then the integrated tail distribution of $F(x)$ (see for example Yang (2004)) is given by $F_I(x) = \frac{1}{E(X)}\int_0^x \bar{F}(y)\,dy$. For the distribution proposed in this work, as is proven in the following result, the integrated tail distribution can be written as a closed-form expression.

Proposition 3. *The integrated tail distribution of the cumulative distribution function $F(x) = 1-\bar{F}(x)$, where $\bar{F}(x)$ is given in (5) results*

$$F_I(x) = 1 + \frac{1-\theta}{\mu\alpha\sigma\theta}\left[\sigma\tan^{-1}\alpha - x\tan^{-1}\left(\alpha\left(\frac{\sigma}{x}\right)^{\theta}\right)\right]$$

$$-\frac{1}{\mu}\left(\frac{\sigma}{x}\right)^{\theta-1}\Psi(\sigma,\theta,\alpha), \qquad (12)$$

where

$$\Psi(\sigma,\theta,\alpha) = {}_2F_1\left(1, \frac{\theta-1}{2\theta}; \frac{3\theta-1}{2\theta}; -\alpha^2\left(\frac{\sigma}{x}\right)^{2\theta}\right),$$

and μ is the mean value of the distribution, given in (11).

Proof. The result follows after some algebra. ■

https://doi.org/10.1017/asb.2015.9 Published online by Cambridge University Press

Under the classical model (see Yang (2004)) and assuming a positive security loading, ρ, for the claim size distributions with regularly varying tails we have that, by using (12), it is possible to obtain an approximation of the probability of ruin, $\Psi(u)$, when $u \to \infty$. In this case, the asymptotic approximations of the ruin function are given by

$$\Psi(u) \sim \frac{1}{\rho}\bar{F}_I(u), \quad u \to \infty,$$

where $\bar{F}_I(u) = 1 - F_I(u)$.

4.3. Mean excess function

The failure rate of the integrated tail distribution, which is given by $\gamma_I(x) = \bar{F}(x)/\int_x^\infty \bar{F}(y)\,dy$, is also obtained in closed-form. Furthermore, the reciprocal of $\gamma_I(x)$ is the mean residual life that can be easily derived.

For a claim amount random variable X, the mean excess function or mean residual life function is the expected payment per claim on a policy with a fixed amount deductible of x, where claims with amounts less than or equal to x are completely ignored:

$$e(x) = E(X - x | X > x) = \frac{1}{\bar{F}(x)} \int_x^\infty \bar{F}(u)\,du.$$

For the PAT distribution proposed here, the mean excess function results

$$e(x) = x\left(\frac{\alpha\theta\left(\frac{\sigma}{x}\right)^\theta\,\Psi(\sigma,\theta,\alpha)}{(\theta-1)\tan^{-1}\left(\alpha\left(\frac{\sigma}{x}\right)^\theta\right)} - 1\right).$$

5. Application to a real data set

In this section, the versatility of (6), as compared with different heavy tail distributions is proven by analyzing real actuarial loss data. This set of data describes a Norwegian fire insurance portfolio from 1989 to 1992 (see Beirlant *et al.* 1996). The Norwegian fire claims data have been recently used by Brazauskas and Kleefeld (2011) where log-folded-normal and log-folded-t families were used to describe total damage caused by 827 fires in Norway for the year 1988 (see also Scollnik (2014) and Brazauskas and Kleefeld (2014)). This data set includes the claim value on 2,585 fire insurance losses in Norwegian Krone ($\times 1000$ NOK). A priority of 500 units was in force, thus no claims below this limit were recorded. For that reason, several probability laws with heavy tail defined in $\mathbb{R}^+$, such as the classical Pareto, Shifted Lognormal, Burr, Loggamma, Fréchet, Inverse Gamma and Stoppa distributions, have been selected to explain the claim amount distribution. Firstly, score equations and second derivatives of the log-likelihood function of the PAT distribution are given; next, it is shown that the distribution introduced in this manuscript outperforms the models mentioned above in terms of different measures of model validation. Finally, some actuarial applications of (6) are given.

5.1. Estimation and model assessment

We have fitted the PAT distribution and all the models described above to the Norwegian fire claim data. Parameters have been estimated by the method of maximum likelihood. In the following, it will be assumed that $\{x_1, x_2, \ldots, x_n\}$ is a random sample selected from the distribution (6). Then, the log-likelihood function is given by

$$\ell \equiv \ell(\alpha, \theta \mid x_1, \ldots, x_n) = -n \log(\tan^{-1}\alpha) + n(\log\alpha + \theta\log\sigma + \log\theta) + (\theta - 1)\sum_{i=1}^{n} \log x_i - \sum_{i=1}^{n} \log\left[x_i^{2\theta} + (\alpha\sigma^{\theta})^2\right].$$

The first derivative of this function with respect to the parameters α and θ, assuming that σ is known, are given by,

$$\frac{\partial\ell}{\partial\alpha} = n\left[\frac{1}{\alpha} - \frac{1}{(1+\alpha^2)\tan^{-1}\alpha}\right] - 2\alpha\sigma^{2\theta}\sum_{i=1}^{n}\frac{1}{x_i^{2\theta} + (\alpha\sigma^{\theta})^2} = 0,$$

$$\frac{\partial\ell}{\partial\theta} = n\left(\log\sigma + \frac{1}{\theta}\right) + \sum_{i=1}^{n}\log x_i - 2\sum_{i=1}^{n}\frac{x_i^{2\theta}\log x_i + (\alpha\sigma^{\theta})^2\log\sigma}{x_i^{2\theta} + (\alpha\sigma^{\theta})^2} = 0.$$

The second partial derivatives, which are used to approximate Fisher's information matrix, are as follows:

$$\frac{\partial^2\ell}{\partial\alpha^2} = n\left[\frac{2\alpha\tan^{-1}\alpha + 1}{(1+\alpha^2)^2(\tan^{-1}\alpha)^2} - \frac{1}{\alpha^2}\right] + 2\sigma^{2\theta}\sum_{i=1}^{n}\frac{(\alpha\sigma^{\theta})^2 - x_i^{2\theta}}{\left[x_i^{2\theta} + (\alpha\sigma^{\theta})^2\right]^2},$$

$$\frac{\partial^2\ell}{\partial\theta^2} = -\frac{n}{\theta^2} - 4\sum_{i=1}^{n}\frac{x_i^{2\theta}(\alpha\sigma^{\theta})^2(\log x_i - \log\sigma)^2}{\left[x_i^{2\theta} + (\alpha\sigma^{\theta})^2\right]^2},$$

$$\frac{\partial^2\ell}{\partial\alpha\partial\theta} = 4\alpha\sigma^{2\theta}\left[\sum_{i=1}^{n}\frac{x_i^{2\theta}\log x_i + (\alpha\sigma^{\theta})^2\log\sigma}{\left[x_i^{2\theta} + (\alpha\sigma^{\theta})^2\right]^2} - \log\sigma\sum_{i=1}^{n}\frac{1}{x_i^{2\theta} + (\alpha\sigma^{\theta})^2}\right].$$

Table 1 provides parameter estimates together with standard errors (in brackets) computed by inverting the approximations of the observed information matrices. Different measures of goodness-of-fit based on information-criterion approach are also given in this table. In this sense, in the third column

TABLE 1

ESTIMATED VALUES OF DIFFERENT HEAVY-TAILED MODELS FOR NORWEGIAN FIRE INSURANCE LOSS DATA.

Distribution	Parameter Estimates (S.E.)	NLL	AIC	BIC	CAIC
Pareto	$\theta = 1.0350$ (0.0204) $\sigma = 500$	21058.50	42119.00	42124.85	42125.85
Shifted Lognormal	$\mu = 6.4307$ (0.0286) $b = 478.2640$ (3.2020) $\sigma = 1.3122$ (0.0244)	20993.60	41993.19	42010.76	42013.76
Stoppa	$\alpha = 1.3536$ (0.0370) $\beta = 1.6434$ (0.0831) $b = 486.4459$ (4.0103)	20984.56	41975.13	41992.70	41995.70
Frećhet	$\alpha = 1.2452$ (0.0414) $s = 559.317$ (22.5823) $b = 325.743$ (15.6615)	21024.36	42054.71	42072.29	42075.29
Inverse Gamma	$\alpha = 1.3419$ (0.0521) $b = 351.739$ (10.1163) $\beta = 746.932$ (51.0971)	21018.43	42042.86	42060.43	42063.43
Loggamma	$\alpha = 1.6607$ (0.0124) $b = 499.369$ (0.1409) $\beta = 3.8067$ (0.0647)	22527.22	45060.45	45078.02	45081.02
PAT	$\alpha = 2.0815$ (0.1371) $\sigma = 500$ $\theta = 1.4890$ (0.0390)	20935.66	41875.31	41887.03	41889.03

the negative of the maximum of the log-likelihood (NLL) is provided. Next in the fourth column Akaike's Information Criteria (AIC, which is calculated by twice NLL plus twice the number of parameters), evaluated at the maximum likelihood estimates, the Bayesian information criterion (BIC, which is obtained as twice the NLL at the estimates plus $k\ln(n)$, where k is the number of free parameters and n is the sample size) and a corrected version of the AIC, Consistent Akaike's Information Criteria (CAIC) were proposed by Bozdogan (1987) to overcome the tendency of the AIC to overestimate the complexity of the underlying model since it lacks certain properties of asymptotic consistency as it does not directly depend on the sample size. Then, in order to calculate the CAIC, a correction factor based on the sample size is used to compensate for the overestimating nature of AIC. The CAIC is defined as twice NLL plus $k(1+\ln(n))$, again k is the number of free parameters and n refers to the sample size.

A lower value of these measures is desirable. These results show that the PAT distribution provides a better fit than do the classical Pareto, shifted Lognormal, Stoppa, Fréchet, Inverse Gamma and Loggamma distributions, even

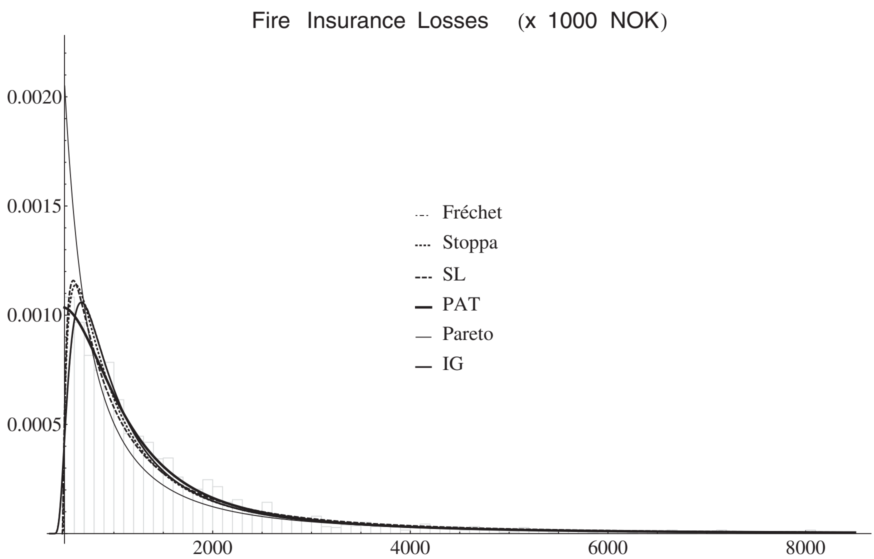

FIGURE 4: Fitted density curves and histogram for Norwegian fire insurance loss data for Pareto (solid thin), Shifted Lognormal (SL, dashed), Stoppa (dotted), Fréchet (dot-dashed), Inverse Gamma (IG, solid) and PAT (solid thick).

when some of these distributions use a larger number of parameters. The Burr distribution has also been fitted to this set of data; however, due to its poor performance (a value of 23670.39 was obtained for NLL), this model has not been included in Table 1. Note that the threshold for the PAT and the classical Pareto is set equal to 500.

As a result of the high values obtained for the different measures of model validation, the Loggamma distribution will be no longer considered in the remainder of this section. The main portion of the density curves for some of the models considered in Table 1 is shown in Figure 4 with an empirical histogram of the data overlaid on top. It can be observed that the PAT distribution best describes the losses.

Now, we analyze model validation from a practical perspective. In this regard, the Pareto distribution can be seen as a limiting case of PAT distribution when α tends to zero. We are interested, by means of the likelihood ratio test, in determining whether the Pareto distribution (null hypothesis) is preferable to the PAT model (alternative hypothesis) to describe this set of data. The test statistic is $T = 2\,(\ell_{\text{PAT}} - \ell_{\text{Pareto}})$ where ℓ_{PAT} and ℓ_{Pareto} represent the maximum of the log-likelihood function for the PAT and Pareto distributions, respectively. Asymptotically, it follows a chi-square distribution with one degree of freedom. Here, it is verified that $T = 2\,(-20935.66 + 21058.50) = 245.687$. Then, at the 5% significance level, as $245.687 > 3.841$ the null hypothesis is clearly rejected and consequently, the smaller model (Pareto) is rejected in favor of the model based on the PAT distribution.

Next, the likelihood ratio test proposed by Vuong (see Vuong (1989)) will be considered as a tool for model diagnostic. The test statistic is

$$T = \frac{1}{\omega\sqrt{n}}\left(\ell_f(\widehat{\theta_1}) - \ell_g(\widehat{\theta_2}) - \log n\left(\frac{p}{2} - \frac{q}{2}\right)\right),$$

where

$$\omega^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\log\left(\frac{f(\widehat{\theta_1})}{g(\widehat{\theta_2})}\right)\right]^2 - \left[\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{f(\widehat{\theta_1})}{g(\widehat{\theta_2})}\right)\right]^2$$

is the estimated variance calculated in the usual manner and f and g represent the probability density function (pdf) of two different non-nested models, $\widehat{\theta_1}$ and $\widehat{\theta_2}$ are the maximum likelihood estimators of θ_1 and θ_2 and p and q are the number of estimated coefficients in the model with pdf f and g, respectively. Note that the Vuong statistic is sensitive to the number of estimated parameters in each model and therefore the test must be corrected for dimensionality. Under the null hypotheses, $H_0 : E[\ell_f(\widehat{\theta_1}) - \ell_g(\widehat{\theta_2})] = 0$, T is asymptotically normally distributed. It is generally accepted that rejection region for this test in favor of the alternative hypothesis occurs, at the 5% significance level, when $T > 1.96$.

Now we compare the Shifted Lognormal and PAT distribution in terms of the Vuong test. Under the null hypothesis, the two models are equally close to the true but unknown specification. In this case, as $T = 51.204$, H_0 is rejected; therefore, differences between these two models exist. Similarly, Stoppa and PAT distribution can also be studied. Here, as $T = 56.952$, the null hypothesis is again rejected and PAT distribution is preferred. Finally, the PAT distribution is preferred to the Fréchet distribution ($T = 50.580$) and Inverse Gamma distribution ($T = 58.779$) at the 5% significance level.

In the following, two more measures of model assessment for individual data based on the empirical distribution function such as the Kolmogorov–Smirnov (KS) test and Crámer von Mises (CvM) test (see Brazauskas and Serfling (2003) and Rizzo (2009) for details) have been applied to these six models. For these tests, smaller values indicate a better fit of the distribution to the data. The results are shown in Table 2. As can be observed for the KS test the PAT model when $\sigma = 500$ outperforms widely the other five distributions. As this test is relatively insensitive to deviations in the tail, the CvM has also been considered to reflect the effect of the tail on the different models. Here, the PAT distribution exceeds Pareto, Shifted Lognormal, Stoppa, Fréchet and Inverse Gamma distributions. These tests also allow us to perform hypothesis testing for model validation purposes. The p-value of the test statistics, computed using via Monte Carlo methods using a simulation size of 10,000 repetitions are also displayed in Table 2. An extremely small p-value may lead to a confident rejection of the null hypothesis that the data comes from the proposed model. As can be seen, the PAT distribution has relatively high p-values across these measures of model assessment and is not rejected and the model is statistically significant.

TABLE 2

GOODNESS-OF-FIT TESTS AND THEIR CORRESPONDING p-VALUES FOR FITTED PARETO, SHIFTED LOGNORMAL, STOPPA, FRÉCHET, INVERSE GAMMA AND PAT FOR NORWEGIAN FIRE INSURANCE LOSS DATA.

	Pareto	Shifted Lognormal	Stoppa	Fréchet	Inverse Gamma	PAT
KS	0.1066	0.0331	0.0403	0.0321	0.0308	0.0124
p-value	0.0000	0.0086	0.0006	0.0096	0.0146	0.8202
CvM	9.9848	0.6532	0.9759	0.7869	0.7144	0.0889
p-value	0.0000	0.0162	0.0018	0.0082	0.0134	0.6340

5.2. Point estimation of high quantiles

High quantiles of the distribution of the claim amounts have been traditionally considered as a measure that provides useful information for practitioners. These quantiles can be estimated by their empirical counterparts. Similarly, let $q_0 = \Pr(X \leq x_0) = F(x_0)$, $0 \leq q \leq 1$, be the values of q_0 given the corresponding x_0. For the PAT distribution, the values of x_0, given by $x_0 = F^{-1}(q_0)$, can be easily computed by taking into account that $x_0 = \sigma \left[\frac{1}{\alpha}\tan((1-q_0)\tan^{-1}\alpha)\right]^{-1/\theta}$. For the sake of comparison, the Log–Log plots for the six models considered above are presented in Figure 5. This chart consists of plots of the logarithm of observed quantiles against the logarithm of theoretical quantiles for each one of the fitted models. As usual, estimated log-quantiles are plotted on the horizontal axis and the logarithm of ordered observations on the vertical axis, where $F^{-1}(q)$ is the estimated q-th quantile and $q = \frac{j}{n+1}$ with $j = 1, \ldots, n$. According to these charts, for high quantiles, Pareto, Stoppa, Fréchet and Inverse Gamma distributions tend to overestimate the observed data whereas the Shifted Lognormal underestimates the empirical data. It can also be observed that the PAT distribution seems to be a reasonable choice for the given data.

Furthermore, empirical and fitted models quantiles for the six models considered in the extreme portion of the tail are exhibited in Table 3. It is of interest to study how close theoretical tail quantiles for each fitted model are from the empirical quantiles. Again, the PAT distribution provides the best fit to data. The Pareto, Stoppa, Fréchet and Inverse Gamma distributions tend to overestimate the extreme tail quantiles whereas the Shifted Lognormal model underestimates them. At this point, it is important to mention that we must be prudent in the conclusions obtained from this table since the sample size is 2,585, and for example, the value of 99.99% empirical quantile represents an event that occurs 1 in 10,000 times.

In Table 4, the tail value at risk (TVaR) for different security levels has been calculated for the models considered. This risk measure describes the expected loss given that the loss exceeds the security level (quantile). These values have been calculated directly from the data. Empirical values have also been

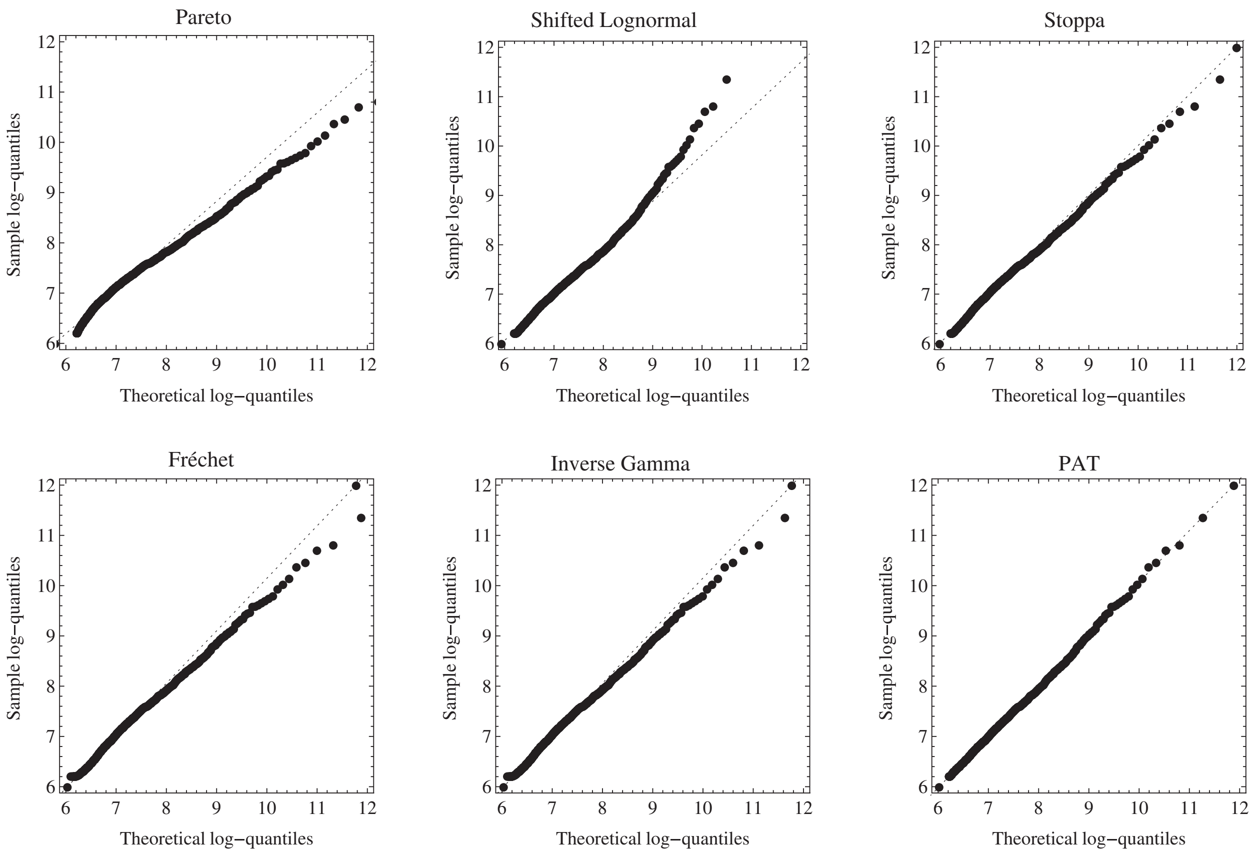

FIGURE 5: Log–Log plots of some of the models considered for Norwegian fire insurance loss data.

https://doi.org/10.1017/asb.2015.9 Published online by Cambridge University Press

TABLE 3

EMPIRICAL AND FITTED QUANTILES.

		Fitted Model					
Quantiles	Empirical	Pareto	Shifted Lognormal	Stoppa	Fréchet	Inverse Gamma	PAT
0.50	1121	976.84	1098.86	1069.81	1076.49	1078.91	1116.93
0.90	3465	4625.61	3813.55	3790.38	3734.1	3655.30	3542.90
0.95	5452	9036.99	5850.68	6373.89	6401.82	6116.85	5655.32
0.99	17150	42792.49	13616.46	21055.39	22823.00	20238.67	16679.65
0.999	85786	395882.52	36275.88	115536.02	143814.53	112422.96	78308.51
0.9995	102438	773430.21	47036.53	192817.18	250743.93	188424.02	124733.70
0.9999	145156	3662394.25	82170.77	633228.53	912518.28	625123.19	367636.86

TABLE 4

TAIL VALUE AT RISK (TVAR) FOR THE DIFFERENT MODELS CONSIDERED.

		Fitted Model					
Security Level	Empirical	Pareto	Shifted Lognormal	Stoppa	Fréchet	Inverse Gamma	PAT
0.90	9936.60	136861.19	7997.18	14796.88	18381.26	14286.18	10814.99
0.95	15635.29	267267.40	11331.77	24728.08	31935.62	23931.78	17231.80
0.99	42475.20	1266000.34	23267.94	81290.92	115755.98	79343.15	50793.41
0.999	166690.00	11712184.71	55815.59	445604.80	734310.50	441154.50	238461.96

obtained. Again, note that the same comment made at the end of the previous paragraph applies here.

5.3. Limited expected value

Finally, in order to choose a model that provides an acceptable description of the loss process, we should verify that $L(x)$ and the empirical limited expected value function which is given by $E_n(x) = \frac{1}{n}\sum_{i=1}^{n} \min(x_i, x)$ are essentially in agreement. Obviously when x tends to infinity, $L(x)$ and $E_n(x)$ approach $E(X)$ and the sample mean, respectively.

In Table 5, empirical and fitted limited expected value functions for the six models are considered. As can be observed, the PAT distribution stays closer to the empirical limited expected value than does the other three models for all the values considered.

TABLE 5

LIMITED EXPECTED VALUE FOR THE DISTRIBUTIONS CONSIDERED AND DIFFERENT VALUES OF THE POLICY LIMIT x.

		Limited expected value function					
Policy limit x ($\times 1,000$ NOK)	Empirical	Pareto	Shifted Lognormal	Stoppa	Fréchet	Inverse Gamma	PAT
1,000	882.61	842.41	876.06	875.44	882.42	882.75	883.12
2,000	1247.97	1176.61	1248.11	1224.73	1226.53	1227.03	1244.48
5,000	1574.90	1606.14	1622.60	1593.49	1579.53	1570.21	1589.04
10,000	1752.60	1922.05	1796.85	1807.14	1791.83	1770.96	1765.67
20,000	1892.30	2230.39	1877.75	1975.28	1968.22	1926.94	1891.59
50,000	2023.13	2626.68	1918.46	2143.22	2157.97	2083.78	2004.32
100,000	2086.70	2918.13	1928.36	2238.71	2279.91	2178.72	2061.65
120,000	2095.38	2993.63	1929.43	2260.18	2305.79	2197.34	2073.77
∞	2105.11	14792.36	1946.09	2582.65	2936.43	2536.40	2203.75

6. CONCLUSIONS

In this paper, a new mechanism to derive probability distributions by adding, from the truncated half-Cauchy distribution, a parameter to a parent distribution function has been presented. Although the method has only been applied to the classical Pareto distribution, certainly this procedure can be extended by allowing the choice of other probabilistic families as parent distribution. Surely, more flexible distributions will be obtained. Furthermore, analysis and applications of these models remain as a topic for further study.

The model proposed here, called Pareto ArcTan (PAT), is a very versatile distribution allowing closed-form expressions for a lot of interesting results in the insurance context. Furthermore, the model seems suitable for modelling payments that include a positive priority, with no claims below that threshold, and losses combine data with high frequencies near the lower limit together with large upper tail derived from massive losses with low frequencies.

The performance of this new family has been illustrated using the well-known Norwegian insurance fire claim data. Numerical results show that the PAT distribution outperforms, in the context of this set of data, other existing long-tail distributions under the different measures of model assessment considered and analysis of high quantiles.

Finally, derivation of composite and folded models based on the PAT distribution might be a line of further research following the work of Nadarajah and Bakar (2014), Scollnik and Sun (2012) and Scollnik (2014), among others. In this sense, it might be interesting to explore the development of spliced models based on Lognormal or Weibull and PAT distributions applying continuity and differentiability conditions at a threshold value.

Acknowledgments

The authors would like to express their gratitude to two anonymous referees for their relevant and useful comments. Authors are funded by grant ECO2013-47092 (Ministerio de Economía y Competitividad, Spain). EGD also acknowledges the Faculty of Business and Economics and the Centre of Actuarial Studies at the University of Melbourne for their special support, since part of this paper was written while EGD was visiting the University of Melbourne in July 2014.

References

ARNOLD, B.C. (1983) *Pareto Distributions*. Silver Spring, MD: International Cooperative Publishing House.

BEIRLANT, J., TEUGELS, J.L AND VYNCKIER, P. (1996) *Practical Analysis of Extreme Values*. Leuven, Belgium: Leuven University Press.

BINGHAM, N.H., GOLDIE, C.M. AND TEUGELS, J.L. (1987) *Encyclopedia of Mathematics and its Applications (No. 27) Regular Variation.* Cambridge: Cambridge University Press.

BOYD, A.V. (1988) Fitting the truncated Pareto distribution to loss distributions. *Journal of the Staple Inn Actuarial Society*, **31**, 151–158.

BOZDOGAN, H. (1987) Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, **52**(3), 345–370.

BRAZAUSKAS, V. and KLEEFELD, A. (2011) Folded and log–folded–*t* distributions as models for insurance loss data. *Scandinavian Actuarial Journal*, 2011(1), 59–74.

BRAZAUSKAS, V. and KLEEFELD, A. (2014) Authors' reply to "Letter to the Editor: Regarding folded models and the paper by Brazauskas and Kleefeld (2011)" by Scollnik. *Scandinavian Actuarial Journal*, **2014**(8), 753–757.

BRAZAUSKAS, V. and SERFLING, R. (2003) Favorable estimator for fitting Pareto models: A study using goodness-of-fit measures with actual data. *Astin Bulletin*, **33**(2), 365–381.

CASTELLANOS, D. (1988) The ubiquitous pi. *Mathematics Magazine*, **61**, 67–98.

JACOB, E. and JAYAKUMAR, K. (2012) On half–Cauchy distribution and process. *International Journal of Statistika and Mathematika*, **3**(2), 77–81.

KLEIBER, C. and KOTZ, S. (2003) *Statistical Size Distributions in Economics and Actuarial Sciences*. Hoboken, NJ: John Wiley & Sons.

KLUGMAN, S.A., PANJER, H.H. and WILLMOT, G.E. (2008) *Loss Models: From Data to Decisions*, 3rd ed. Wiley.

MARSHALL, A.W. and OLKIN, I. (1997) A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, **84**(3), 641–652.

MATA, A. (2000) Princing excess of loss reinsurance with reinstatements. *Astin Bulletin*, **30**(2), 349–368.

NADARAJAH, S. and BAKAR, S.A.A. (2014) New composite models for the Danish fire insurance data. *Scandinavian Actuarial Journal*, **2**, 180–187.

RAJAN, S., WANG, S., INKOL, R. and JOYAL, A. (2006) Efficient approximations for the arctangent function. *IEEE Signal Processing Magazine*, **23**(3), 108–111.

RIZZO, M.L. (2009) New goodness-of-fit tests for Pareto distributions. *Astin Bulletin*, **39**(2), 691–715.

ROLSKI, T., SCHMIDLI, H. SCHMIDT, V. and TEUGEL, J. (1999) *Stochastic Processes for Insurance and Finance*. John Wiley & Sons.

RYTGAARD, M. (1990) Estimation in the Pareto distribution. *Astin Bulletin*, **20**(2), 201–216.

SARABIA, J.M. and CASTILLO, E. (2005) About a class of max–stable families with applications to income distributions. *Metron*, LXIII, **3**, 505–527.

SCOLLNIK, D.P.M. (2014) Regarding folded models and the paper by Brazauskas and Kleefeld (2011) *Scandinavian Actuarial Journal*, **2014**(3), 278–281.

SCOLLNIK, D.P.M. and SUN, C. (2012) Modeling with Weibull–Pareto models. *North American Actuarial Journal*, **16**(2), 260–272.

VUONG, Q. (1989) Likelihood ratio tests for model selection and non-nested hypotheis. *Econometrica*, **50**, 1–25.

YANG, H. (2004) Cramér–Lundberg asymptotics. In *Encyclopedia of Actuarial Science*. Wiley.

EMILIO GÓMEZ–DÉNIZ (Corresponding author)
Department of Quantitative Methods in Economics and TiDES Institute.
University of Las Palmas de Gran Canaria, Spain
E-Mail: emilio.gomez-deniz@ulpgc.es

ENRIQUE CALDERÍN–OJEDA
Centre for Actuarial Studies, Department of Economics, The University of Melbourne, Australia
E-Mail: ecalderin@unimelb.edu.au