

A new family of tempered distributions

Lucio Barabesi

*Department of Economics and Statistics, University of Siena
Piazza S. Francesco 7, 53100 Siena, Italy
e-mail: lucio.barabesi@unisi.it*

Andrea Cerasa

*European Commission, Joint Research Centre
Competences Directorate, Text and Data Mining Unit
Via E. Fermi, 21027 Ispra, Italy
e-mail: andrea.cerasa@jrc.ec.europa.eu*

Andrea Cerioli

*Department of Economics and Management, University of Parma
Via Kennedy 6, 43125 Parma, Italy
e-mail: andrea.cerioli@unipr.it*

and

Domenico Perrotta

*European Commission, Joint Research Centre
Competences Directorate, Text and Data Mining Unit
Via E. Fermi, 21027 Ispra, Italy
e-mail: domenico.perrotta@ec.europa.eu*

Abstract: Tempered distributions have received considerable attention, both from a theoretical point of view and in several important application fields. The most popular choice is perhaps the Tweedie model, which is obtained by tempering the Positive Stable distribution. Through tempering, we suggest a very flexible four-parameter family of distributions that generalizes the Tweedie model and that could be applied to data sets of non-negative observations with complex (and difficult to accommodate) features. We derive the main theoretical properties of our proposal, through which we show its wide application potential. We also embed our proposal within the theory of Lévy processes, thus providing a strengthened probabilistic motivation for its introduction. Furthermore, we derive a series expansion for the probability density function which allows us to develop algorithms for fitting the distribution to data. We finally provide applications to challenging real-world examples taken from international trade.

MSC 2010 subject classifications: Primary 62E15, 60E07; secondary 62P20.

Keywords and phrases: Heavy tails and point mass at zero, international trade, Positive Linnik distribution, subordination of Lévy processes, tempering, Tweedie distribution.

Received July 2016.

Contents

1	Introduction	3872
2	Two distribution families related to the Positive Stable law	3873
2.1	The Positive Linnik distribution	3873
2.2	The Tweedie distribution	3877
3	The new family of distributions	3878
4	Further properties of the new family	3880
5	Connection with Lévy processes	3882
6	Application to international trade	3886
	Acknowledgements	3891
	References	3891

1. Introduction

There already exist rather flexible and probabilistically founded models that have experienced successful applications in different areas. For instance, a popular class is provided by stable distributions and by their tempered versions. Such models possess a nice probabilistic interpretation in terms of Lévy processes and have proven to be useful for describing phenomena that exhibit skewness and kurtosis, especially in economics and finance; see, e.g., Barndorff-Nielsen and Shephard [9], Brix [10], Carr et al. [11], Cont and Tankov [15], Devroye and James [16], Favaro et al. [19, 20], Lijoi and Prunster [33], Rachev et al. [39].

It is worth noting that even if stable distributions represent a very attractive mathematical framework, they may not provide realistic descriptions in many practical applications. Indeed, stable distributions do not generally display finite moments, while the size of random phenomena are bounded in real-world situations. This shortcoming has motivated the study of their tempered versions (Grabchak [23]). Tempering allows for models that are similar to original distributions in some central region, even if they possess lighter – i.e., tempered – tails. Therefore, even if the original stable distribution and its tempered version may be statistically indistinguishable, their tail behavior is different in the sense that the former may have infinite expectation, while the latter possesses all finite moments.

A frequently adopted choice for tempering the Positive Stable distribution is the three-parameter Tweedie model (Jørgensen [28], Tweedie [43]), which encompasses both the exponentially-tilted stable distribution and the compound Poisson of Gamma distributions as special cases. The possibility to represent skewness, heavy tails and a point mass at zero (the latter arising, e.g., as a consequence of rounding errors or confidentiality issues in real surveys) all under the same umbrella has made the Tweedie distribution a very attractive framework for modelling international trade data, where such features occur very frequently (Barabesi et al. [5]). However, a possible limitation of the Tweedie model is that it may not adequately describe data that simultaneously exhibit

both a heavy tail and the point mass at zero. It is not a fault of this three-parameter system, which can control scale, shape and tail heaviness through its parameters, but which might have problems in accommodating both skewness and kurtosis simultaneously. As also shown by Barabesi et al. [5], alternative and rather ad hoc three-parameter mixtures – that might occasionally provide acceptable results – lack a general probabilistic interpretation and show highly variable performance over different data sets.

The goal of this paper is to introduce a very flexible four-parameter family of distributions that overcomes the shortcomings described above and that could be applied to data sets of non-negative observations with complex (and difficult to accommodate) features, thus improving over the Tweedie model. For this purpose we rely on the three-parameter Positive Linnik distribution (Pakes [37]), which provides an extension to the Positive Stable law. The proposed family is obtained by tempering the Positive Linnik distribution and yields the Tweedie distribution in the limit as the shape parameter approaches zero. We derive the main theoretical properties of our four-parameter system, including various stochastic representations, through which we show its wide application potential. We also embed our proposal within the theory of Lévy processes, thus providing a strengthened probabilistic motivation for its introduction. Furthermore, we derive explicit expressions of the probability density function in terms of the Mittag-Leffler function. These expressions allow us to develop algorithms for fitting the distribution to data. In addition, random variate generation is considered by means of the achieved stochastic representations. A related proposal for genuinely integer-valued data is given in Barabesi et al. [3].

The paper is organized as follows. In §2 we review the Positive Linnik distribution and the Tweedie model. Our new family of distributions is proposed in §3. Its main theoretical properties are derived in §4, while §5 provides connections with the theory of Lévy processes. Applications to challenging real-world examples taken from international trade are provided in §6.

2. Two distribution families related to the Positive Stable law

Before introducing the new distribution, we describe and provide some basic features of two well-known laws, that are connected to our proposal. These are the Positive Linnik and the Tweedie families. Both of them encompass the Positive Stable law either as a special case, or in the limit. A repeatedly-adopted notation in our work is as follows. If X represents a random variable, the corresponding Laplace transform is given by $L_X(s) = E[\exp(-sX)]$, $\text{Re}(s) > 0$. In addition, if X is an integer-valued random variable, the corresponding probability generating function (p.g.f.) is given by $g_X(s) = E[s^X]$, $s \in [0, 1]$.

2.1. The Positive Linnik distribution

The Positive Linnik (PL) law was originally given by Pakes [37] and extends the law introduced by Linnik [35, p.67]. The law is also named Generalized Mittag-Leffler, since it involves the generalized Mittag-Leffler function (see Haubold

et al. [24]). By introducing a slight change in the parametrization suggested by Christoph and Schreiber [14], the absolutely continuous PL random variable, say X_{PL} , has Laplace transform given by

$$L_{X_{PL}}(s) = (1 + \lambda \delta s^\gamma)^{-1/\delta}, \quad \operatorname{Re}(s) > 0, \quad (1)$$

where $(\gamma, \lambda, \delta) \in]0, 1] \times \mathbb{R}^+ \times \mathbb{R}^+$, with $\mathbb{R}^+ =]0, \infty[$. It is worth noting that γ is a “tail” index (see below for more details), while λ is actually the scale parameter and δ is the shape parameter. We also adopt the notation $X_{PL} := X_{PL}(\gamma, \lambda, \delta)$ to emphasize dependence on the parameters.

The PL family encompasses the Positive Mittag-Leffler law (originally proposed by Pillai [38]) for $\delta = 1$, and the Gamma law for $\gamma = 1$. In addition, the Laplace transform of the Positive Stable (PS) random variable $X_{PS} := X_{PS}(\gamma, \lambda)$, i.e.

$$L_{X_{PS}}(s) = \exp(-\lambda s^\gamma), \quad \operatorname{Re}(s) > 0,$$

is obtained as a special case of (1) when $\delta \rightarrow 0^+$. Obviously, the law of X_{PS} reduces to a Dirac mass at λ for $\gamma = 1$. See Zolotarev [45] for more details on the PS law.

We now recall a key property of the PL law. Let the random variable $X_G := X_G(\alpha, \beta)$ be distributed according to a Gamma law with Laplace transform given by

$$L_{X_G}(s) = (1 + \alpha s)^{-\beta}, \quad \operatorname{Re}(s) > 0,$$

where $(\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$. Therefore, the corresponding probability density function (p.d.f.) with respect to the Lebesgue measure on \mathbb{R} is given by

$$f_{X_G}(x) = \frac{1}{\alpha^\beta \Gamma(\beta)} x^{\beta-1} \exp(-x/\alpha), \quad x \in \mathbb{R}^+.$$

Hence, by considering the Gamma random variable $X_G(\lambda\delta, 1/\delta)$, it follows that (1) may be also expressed as

$$L_{X_{PL}}(s) = L_{X_G}(s^\gamma) = \mathbb{E}[\exp(-s^\gamma X_G)],$$

so that the PL distribution is actually a scale mixture of the PS law over a Gamma mixing. The genesis of the PL distribution is thus apparent, since from the previous expression it turns out that (see, e.g., James [26] and Jose et al. [29]),

$$X_{PL}(\gamma, \lambda, \delta) \stackrel{\mathcal{L}}{=} X_G(\lambda\delta, 1/\delta)^{1/\gamma} X_{PS}(\gamma, 1),$$

which may also be rewritten as

$$X_{PL}(\gamma, \lambda, \delta) \stackrel{\mathcal{L}}{=} X_{PS}(\gamma, X_G(\lambda\delta, 1/\delta)). \quad (2)$$

Our notation in (2) actually emphasizes that the conditional distribution of X_{PS} given X_G is PS and the unconditional distribution is PL. In what follows, similar notations have the same meaning. Identity (2) also proves to be useful for random variate generation, since many generators for Gamma variates are

available, while PS variates can be obtained by means of Kanter’s representation (Kanter [30]). The reader may refer to James [26] and Lin [34] for further noteworthy stochastic representations of X_{PL} .

In order to achieve an explicit expression for the p.d.f. of X_{PL} , we make use of the generalized Mittag-Leffler function given by

$$E_{a,b}^c(z) = \sum_{k=0}^{\infty} \frac{\binom{-c}{k}}{\Gamma(ak + b)} (-z)^k,$$

where $\text{Re}(a) \in \mathbb{R}^+$, $\text{Re}(b) \in \mathbb{R}^+$, $c \in \mathbb{R}^+$ and $z \in \mathbb{C}$, while

$$\binom{c}{k} = \frac{c(c-1)\cdots(c-k+1)}{k!}$$

for $c \in \mathbb{R}$ (see, e.g., Mathai and Haubold [36], or the recent monograph by Gorenflo et al. [22]). Reasoning as in Haubold et al. [24, p.37], the p.d.f. of X_{PL} can be obtained by initially expanding (1) as

$$L_{X_{PL}}(s) = (\lambda\delta s^\gamma)^{-1/\delta} (1 + (\lambda\delta s^\gamma)^{-1})^{-1/\delta} = \sum_{k=0}^{\infty} \binom{-1/\delta}{k} (\lambda\delta s^\gamma)^{-k-1/\delta}.$$

Since for $q \in \mathbb{R}^+$ and $\alpha \in \mathbb{R}^+$ it holds that

$$q^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} \exp(-qx) dx,$$

it also follows that

$$\begin{aligned} L_{X_{PL}}(s) &= \sum_{k=0}^{\infty} \binom{-1/\delta}{k} (\lambda\delta)^{-k-1/\delta} \frac{1}{\Gamma(\gamma k + \gamma/\delta)} \int_0^\infty x^{\gamma k + \gamma/\delta - 1} \exp(-sx) dx \\ &= \int_0^\infty \exp(-sx) \frac{1}{x} \sum_{k=0}^{\infty} \frac{\binom{-1/\delta}{k}}{\Gamma(\gamma k + \gamma/\delta)} \left(\frac{x^\gamma}{\lambda\delta}\right)^{k+1/\delta} dx. \end{aligned}$$

Hence, from the uniqueness of Laplace transform, it follows that the p.d.f. of X_{PL} with respect to the Lebesgue measure on \mathbb{R} is given by

$$f_{X_{PL}}(x) = \frac{1}{x} \sum_{k=0}^{\infty} \frac{\binom{-1/\delta}{k}}{\Gamma(\gamma k + \gamma/\delta)} \left(\frac{x^\gamma}{\lambda\delta}\right)^{k+1/\delta} = \frac{x^{\gamma/\delta-1}}{(\lambda\delta)^{1/\delta}} E_{\gamma, \gamma/\delta}^{1/\delta} \left(-\frac{x^\gamma}{\lambda\delta}\right), \quad x \in \mathbb{R}^+. \tag{3}$$

The generalized Mittag-Leffler function may be expressed in terms of the Fox H -function (see, e.g., Haubold et al. [24, p.17]) and it can be computed by means of routines which are commonly available in mathematical software packages, thus allowing practical evaluation of $f_{X_{PL}}$. In any case, it should be noticed that James [26, Remark 3.2] provides an integral representation for $f_{X_{PL}}$. Under our parametrization, this alternative expression of $f_{X_{PL}}$ turns out to be

$$f_{X_{PL}}(x) = \frac{1}{\pi(\lambda\delta)^{1/\gamma}} \int_0^\infty \frac{\exp(-xy/(\lambda\delta)^{1/\gamma}) \sin(\pi\gamma F_\gamma(y)/\delta)}{(y^{2\gamma} + 2y^\gamma \cos(\pi\gamma) + 1)^{1/(2\delta)}} dy, \quad x \in \mathbb{R}^+,$$

where

$$F_\gamma(y) = 1 - \frac{1}{\pi\gamma} \cot^{-1} \left(\cot(\pi\gamma) + \frac{y^\gamma}{\sin(\pi\gamma)} \right), \quad y \in \mathbb{R}^+.$$

If $F_{X_{PL}}$ represents the distribution function (d.f.) of X_{PL} , it is interesting to assess the asymptotic behavior of $F_{X_{PL}}(x)$ as $x \rightarrow \infty$. Since for each $s \in \mathbb{R}^+$ it holds that

$$\frac{L_{X_{PL}}(s\tau)}{L_{X_{PL}}(\tau)} \rightarrow 1$$

as $\tau \rightarrow 0^+$, an application of the Tauberian Theorem (see Feller [21, p.443]) shows that for each $y \in \mathbb{R}^+$

$$\frac{F_{X_{PL}}(xy)}{F_{X_{PL}}(x)} \rightarrow 1$$

as $x \rightarrow \infty$, i.e. $F_{X_{PL}}$ is slowly varying at infinity. It thus follows that

$$F_{X_{PL}}(x) \sim L_{X_{PL}}(1/x) = (1 + \lambda\delta x^{-\gamma})^{-1/\delta}$$

as $x \rightarrow \infty$. In addition, on the basis of the result provided by Feller [21, Example (c), p.447], it also holds that

$$1 - F_{X_{PL}}(x) \sim 1 - L_{X_{PL}}(1/x) \sim \lambda x^{-\gamma}$$

as $x \rightarrow \infty$, and it is now apparent that γ is a “tail” index. Moreover, by considering the previous expression, $F_{X_{PL}}$ interestingly turns out to be asymptotically equivalent to the d.f. of a Burr III distribution, also known as the Dagum distribution, a popular and flexible model for income and economic data; see Fattorini and Lemmi [18] and Kleiber [32] for the genesis of this law.

As Christoph and Schreiber [14] emphasize, a further property of the PL law is that X_{PL} is self-decomposable and hence infinitely divisible. This follows from representation theorems for self-decomposable random variables in terms of their Laplace transform (see, e.g., Steutel and van Harn [42]).

Finally, by generalizing (2), it is worth noting that we can obtain a “scale” mixture of PS random variables, say X_{MPS} , with a mixing absolutely continuous and positive random variable V , on the basis of the identity in distribution

$$X_{MPS} \stackrel{\mathcal{L}}{=} X_{PS}(\gamma, V). \quad (4)$$

Obviously, the PS law is achieved by assuming a degenerate distribution for V , i.e. $P(V = \lambda) = 1$. The Laplace transform of X_{MPS} turns out to be

$$L_{X_{MPS}}(s) = L_V(s^\gamma), \quad \text{Re}(s) > 0, \quad (5)$$

where L_V is the Laplace transform of V . Families of mixtures of PS random variables can thus be generated on the basis of (4) and (5) by suitably selecting V . As an example, if V is distributed according the Geometric Gamma law, we obtain the Geometric Generalized Mittag-Leffler distribution proposed by Jose et al. [29].

2.2. The Tweedie distribution

The Tweedie distribution is actually a Tempered Positive Stable (TPS) distribution, as pointed out by Hougaard [25]; see also Aalen [1]. We thus denote the Tweedie random variable as X_{TPS} . With a slight change in the parametrization proposed by Hougaard [25] and Aalen [1], the Laplace transform of X_{TPS} is given by

$$L_{X_{TPS}}(s) = \exp(\operatorname{sgn}(\gamma)\lambda(\theta^\gamma - (\theta + s)^\gamma)), \quad \operatorname{Re}(s) > 0, \quad (6)$$

where $(\gamma, \lambda, \theta) \in \{] - \infty, 1] \times \mathbb{R}^+ \times \mathbb{R}^+ \} \cup \{]0, 1] \times \mathbb{R}^+ \times \{0\} \}$. Formulation (6) does not require to define the Laplace transform for analytical continuity when $\gamma = 0$. Indeed, under this parametrization it is apparent that the Dirac distribution with mass at zero is obtained for $\gamma = 0$. By following the usual route, we also write $X_{TPS} := X_{TPS}(\gamma, \lambda, \theta)$.

Obviously, $X_{TPS}(\gamma, \lambda, 0) \stackrel{\mathcal{L}}{=} X_{PS}(\gamma, \lambda)$ when $\gamma \in]0, 1]$. Moreover, for $\gamma \in]0, 1]$ it holds that

$$L_{X_{TPS}}(s) = \frac{L_{X_{PS}}(\theta + s)}{L_{X_{PS}}(\theta)}. \quad (7)$$

In such a case, if $f_{X_{PS}}$ denotes the p.d.f. of $X_{PS}(\gamma, \lambda)$ with respect to the Lebesgue measure on \mathbb{R} , the p.d.f. of $X_{TPS}(\gamma, \lambda, \theta)$ is obtained from (7) as

$$f_{X_{TPS}}(x) = \frac{\exp(-\theta x) f_{X_{PS}}(x)}{L_{X_{PS}}(\theta)}. \quad (8)$$

Hence, expression (8) reveals the exponential nature of the tempering and it is apparent that θ is actually the tempering parameter.

Since $f_{X_{PS}}$ may be expressed as the convergent series (see, e.g., Sato [41, p.88])

$$f_{X_{PS}}(x) = \frac{1}{x} \sum_{k=1}^{\infty} \frac{1}{k! \Gamma(-k\gamma)} \left(-\frac{x^\gamma}{\lambda} \right)^{-k}, \quad x \in \mathbb{R}^+,$$

when $\gamma \in]0, 1]$, it also follows from (8) that

$$f_{X_{TPS}}(x) = \frac{\exp(-\theta x + \lambda \theta^\gamma)}{x} \sum_{k=1}^{\infty} \frac{1}{k! \Gamma(-k\gamma)} \left(-\frac{x^\gamma}{\lambda} \right)^{-k}, \quad x \in \mathbb{R}^+.$$

It is worth noting that tempering extends the range of parameter values with respect to the PS distribution. Indeed, γ may now assume negative values, even if θ must be strictly positive in such a case. This is an interesting feature, since when $\gamma \in \mathbb{R}^-$, with $\mathbb{R}^- =] - \infty, 0[$, it is possible to reformulate X_{TPS} as a compound Poisson of Gamma random variables. Let $X_P := X_P(\lambda)$ be a Poisson random variable with parameter λ , whose probability generating function (p.g.f.) is given by

$$g_{X_P}(s) = \exp(-\lambda(1 - s)), \quad s \in [0, 1].$$

Thus, by considering a Gamma random variable $X_G(1/\theta, -\gamma)$, expression (6) may be rewritten as

$$L_{X_{TPS}}(s) = \exp(-\lambda\theta^\gamma(1 - L_{X_G}(s))) = E[L_{X_G}(s)^{X_P}],$$

where in this case $X_P = X_P(\lambda\theta^\gamma)$. Since the sum of independent Gamma random variables is in turn a Gamma random variable, from the previous expression the following identity in distribution holds for $\gamma \in \mathbb{R}^-$

$$X_{TPS}(\gamma, \lambda, \theta) \stackrel{\mathcal{L}}{=} X_G(1/\theta, -\gamma X_P(\lambda\theta^\gamma)). \quad (9)$$

Hence, X_{TPS} displays a mixed distribution, given by a convex combination of a Dirac distribution with mass at zero and an absolutely continuous distribution, which proves to be a very useful property for modelling data with an excess of zeroes (Barabesi et al. [5]). On the basis of expression (9), it can be shown that

$$P(X_{TPS} = 0) = \exp(-\lambda\theta^\gamma),$$

while the p.d.f. of X_{TPL} conditional on the event $\{X_{TPL} > 0\}$ is given by

$$h_{X_{TPS}}(x) = \frac{1}{1 - \exp(-\lambda\theta^\gamma)} \frac{\exp(-\theta x - \lambda\theta^\gamma)}{x} \sum_{k=1}^{\infty} \frac{1}{k! \Gamma(-k\gamma)} \left(-\frac{x^\gamma}{\lambda}\right)^{-k},$$

for $x \in \mathbb{R}^+$.

For further features of the Tweedie distribution see, e.g., Barabesi et al. [5], who also describe computationally efficient algorithms for parameter estimation and for random variate simulation.

3. The new family of distributions

By following one of the general paths suggested by Klebanov and Slámová [31] and extending expression (2), we introduce a tempered version of the PL random variable, say $X_{TPL} := X_{TPL}(\gamma, \lambda, \delta, \theta)$, on the basis of the identity in distribution

$$X_{TPL}(\gamma, \lambda, \delta, \theta) \stackrel{\mathcal{L}}{=} X_{TPS}(\gamma, X_G(\lambda\delta, 1/\delta), \theta). \quad (10)$$

The new family is thus defined as the scale mixture of a TPS law over a Gamma mixing. On the basis of (6), Identity (10) yields the Laplace transform of X_{TPL} as

$$L_{X_{TPL}}(s) = (1 + \operatorname{sgn}(\gamma)\lambda\delta((\theta + s)^\gamma - \theta^\gamma))^{-1/\delta}, \quad \operatorname{Re}(s) > 0, \quad (11)$$

where $(\gamma, \lambda, \theta, \delta) \in \{]-\infty, 1] \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \} \cup \{]0, 1] \times \mathbb{R}^+ \times \{0\} \times \mathbb{R}^+ \}$. The Laplace transform of $X_{TPS}(\gamma, \lambda, \delta)$ is obtained from (11) as $\delta \rightarrow 0^+$. Expression (10) also provides a suitable tool for generating TPL variates through the generators available for X_{TPS} (see Barabesi et al. [5]). Alternatively, variate generation could be directly implemented by means of the method proposed by Barabesi and Pratelli [6, 7].

Some comments on the TPL distribution are in order according to the two main subsets of the γ domain that strongly characterize this law. First, let us consider $\gamma \in]0, 1]$. In this case, if $\lambda\delta\theta^\gamma = 1$ expression (11) reduces to

$$L_{X_{TPL}}(s) = (1 + (\lambda\delta)^{1/\gamma} s)^{-\gamma/\delta}, \quad \text{Re}(s) > 0, \tag{12}$$

which is the Laplace transform of a Gamma random variable $X_G((\lambda\delta)^{1/\gamma}, \gamma/\delta)$. On the other hand, if $\lambda\delta\theta^\gamma \neq 1$ and by considering the PL random variable $X_{PL}(\gamma, \psi, \delta)$, for the TPL random variable $X_{TPL}(\gamma, \lambda, \delta, \theta)$ it holds

$$L_{X_{TPL}}(s) = \frac{L_{X_{PL}}(\theta + s)}{L_{X_{PL}}(\theta)}, \tag{13}$$

where $\lambda = \psi/(1 + \psi\delta\theta^\gamma)$. Thus, if $f_{X_{PL}}$ is the p.d.f. of $X_{PL}(\gamma, \psi, \delta)$ with respect to the Lebesgue measure on \mathbb{R} , the p.d.f. of $X_{TPL}(\gamma, \lambda, \delta, \theta)$ is obtained from (13) as

$$f_{X_{TPL}}(x) = \frac{\exp(-\theta x) f_{X_{PL}}(x)}{L_{X_{PL}}(\theta)}. \tag{14}$$

Therefore, an exponentially-tilted Linnik occurs and the exponential nature of the proposed tempering is transparent, as in the case of the Tweedie law.

A stochastic representation generalizing (9) is obtained for $\gamma \in \mathbb{R}^-$. To this aim, let $X_{NB} := X_{NB}(\pi, \kappa)$ be a Negative Binomial random variable with p.g.f. given by

$$g_{X_{NB}}(s) = \left(\frac{1 - \pi}{1 - \pi s} \right)^\kappa = (1 + \phi(1 - s))^{-\kappa}, \quad s \in [0, 1],$$

where $(\pi, \kappa) \in]0, 1[\times \mathbb{R}^+$ and $\phi = \pi/(1 - \pi) \in \mathbb{R}^+$. By considering the random variable $X_G(1/\theta, -\gamma)$, the Laplace transform (11) can be rewritten as

$$L_{X_{TPL}}(s) = (1 + \lambda\delta\theta^\gamma(1 - L_{X_G}(s)))^{-1/\delta}.$$

Therefore, the following identity in distribution holds for $\gamma \in \mathbb{R}^-$

$$\begin{aligned} X_{TPL}(\gamma, \lambda, \theta, \delta) &\stackrel{\mathcal{L}}{=} X_G(1/\theta, -\gamma X_{NB}(\lambda\delta\theta^\gamma/(1 + \lambda\delta\theta^\gamma), 1/\delta)) \\ &\stackrel{\mathcal{L}}{=} X_G(1/\theta, -\gamma X_P(X_G(\lambda\delta\theta^\gamma, 1/\delta))), \end{aligned} \tag{15}$$

where the second identity follows from the well-known representation of the Negative Binomial law as a compound Poisson law (see, e.g., Johnson et al. [27, p.212]). It is apparent from (15) that the TPL random variable is a compound Negative Binomial of Gamma random variables. As a consequence, similarly to the TPS law, X_{TPL} displays a mixed distribution given by a convex combination of a Dirac distribution with mass at zero and an absolutely continuous distribution.

We also emphasize that X_{TPL} is self-decomposable when $\gamma \in]0, 1]$. Indeed, since Identity (13) holds in this case, self-decomposability – and hence infinite divisibility – can be proven by means of Proposition V.2.14 of Steutel and van

Harn [42] and by reminding that X_{PL} is in turn self-decomposable. In contrast, X_{TPL} is not self-decomposable when $\gamma \in \mathbb{R}^-$, since a self-decomposable random variable must be necessarily absolutely continuous (see the discussion in Steutel and van Harn [42]). However, from the results given in §5, it follows that X_{TPL} is infinitely divisible for $\gamma \in \mathbb{R}^-$.

Finally, as in (4) and (5), Identity (10) can be extended to a general “scale” mixture of TPS random variables, say X_{MTPS} , with a mixing absolutely continuous and positive random variable V having Laplace transform L_V . In this case, we consider the following identity in distribution

$$X_{MTPS} \stackrel{\mathcal{L}}{=} X_{TPS}(\gamma, V, \theta). \quad (16)$$

The Tweedie law is obtained by assuming a degenerate distribution for V , i.e. $P(V = \lambda) = 1$. Moreover, it is apparent from (16) that the Laplace transform of X_{MTPS} turns out to be

$$L_{X_{MTPS}}(s) = L_V(\text{sgn}(\gamma)((\theta + s)^\gamma - \theta^\gamma)), \quad \text{Re}(s) > 0. \quad (17)$$

Families of mixtures of TPS random variables can thus be generated on the basis of (16) and (17) by suitably selecting V . As an example, if V is distributed according the Geometric Gamma law, a tempered version of the Geometric Generalized Mittag-Leffler distribution is achieved.

4. Further properties of the new family

We now obtain the p.d.f. of the proposed random variable X_{TPL} according to the two main subsets of the γ domain. We start by considering the case $\gamma \in]0, 1]$. In such a setting, we have to consider two cases, i.e. $\lambda\delta\theta^\gamma = 1$ and $\lambda\delta\theta^\gamma \neq 1$. When $\lambda\delta\theta^\gamma = 1$, from expression (12) it holds that the p.d.f. of X_{TPL} reduces to

$$f_{X_{TPL}}(x) = \frac{1}{(\lambda\delta)^{1/\delta}\Gamma(\gamma/\delta)} x^{\gamma/\delta-1} \exp(-x/(\lambda\delta)^{1/\gamma}), \quad x \in \mathbb{R}^+. \quad (18)$$

On the other hand, if $\lambda\delta\theta^\gamma \neq 1$, from expressions (3) and (14) it turns out that

$$f_{X_{TPL}}(x) = \frac{\exp(-\theta x)x^{\gamma/\delta-1}}{(\lambda\delta)^{1/\delta}} E_{\gamma, \gamma/\delta}^{1/\delta} \left(-\frac{(1 - \lambda\delta\theta^\gamma)x^\gamma}{\lambda\delta} \right), \quad x \in \mathbb{R}^+, \quad (19)$$

which obviously reduces to (3) when $\theta = 0$.

We now address the case $\gamma \in]-\infty, 0]$. By assuming that $\varrho = 1 + \lambda\delta\theta^\gamma$ for the sake of notational simplicity, it follows from (15) that

$$P(X_{TPL} = 0) = \varrho^{-1/\delta}. \quad (20)$$

Moreover, since it holds that

$$\binom{1/\delta + k - 1}{k} = (-1)^k \binom{-1/\delta}{k},$$

the p.d.f. of the random variable X_{TPL} conditional on the event $\{X_{TPL} > 0\}$ is given by

$$\begin{aligned} h_{X_{TPL}}(x) &= \frac{1}{1 - \varrho^{-1/\delta}} \sum_{k=1}^{\infty} \frac{\theta^{-\gamma k}}{\Gamma(-\gamma k)} x^{-\gamma k - 1} \exp(-\theta x) \times \\ &\quad \times \binom{1/\delta + k - 1}{k} \varrho^{-1/\delta} \left(\frac{\varrho - 1}{\varrho}\right)^k \\ &= \frac{1}{\varrho^{1/\delta} - 1} \frac{\exp(-\theta x)}{x} \sum_{k=1}^{\infty} \frac{\binom{-1/\delta}{k}}{\Gamma(-\gamma k)} \left(-\frac{\lambda \delta}{\varrho x^\gamma}\right)^k, \quad x \in \mathbb{R}^+. \end{aligned}$$

In addition, since it holds that

$$\frac{\binom{-1/\delta}{k}}{\Gamma(-\gamma k)} = \frac{(\gamma k - \gamma + \gamma/\delta) \binom{-1/\delta}{k-1}}{\Gamma(-\gamma k + 1)},$$

and by re-indexing the summation, $h_{X_{TPL}}$ may be finally rewritten as

$$\begin{aligned} h_{X_{TPL}}(x) &= \frac{1}{\varrho^{1/\delta} - 1} \frac{\exp(-\theta x)}{x} \sum_{k=0}^{\infty} \frac{(\gamma k + \gamma/\delta) \binom{-1/\delta}{k}}{\Gamma(-\gamma k - \gamma + 1)} \left(-\frac{\lambda \delta}{\varrho x^\gamma}\right)^{k+1} \quad (21) \\ &= \frac{\lambda \delta}{\varrho(\varrho^{1/\delta} - 1)} \frac{\exp(-\theta x)}{x^{\gamma+1}} \times \\ &\quad \times \left\{ E_{-\gamma, -\gamma}^{1/\delta} \left(\frac{\lambda \delta}{\varrho x^\gamma}\right) + (\gamma - \gamma/\delta) E_{-\gamma, -\gamma+1}^{1/\delta} \left(\frac{\lambda \delta}{\varrho x^\gamma}\right) \right\}, \end{aligned}$$

where $x \in \mathbb{R}^+$. As remarked in §2.1, the generalized Mittag-Leffler function can be computed by means of mathematical software packages in order to evaluate (19) and (21).

Finally, by differentiating (11) and equating to zero, it is easily checked that the expectation and the variance of X_{TPL} are given, respectively, by

$$\mu = E[X_{TPL}] = \operatorname{sgn}(\gamma) \gamma \lambda \theta^{\gamma-1} \quad (22)$$

and

$$\sigma^2 = \operatorname{Var}[X_{TPL}] = \delta \mu^2 + \frac{1 - \gamma}{\theta} \mu. \quad (23)$$

It is worth noting that μ does not depend on δ , so that the TPL and the Tweedie random variables have the same expectation if the same values of the parameters γ , λ and θ are chosen. However, since the dispersion index is given by

$$D = \frac{\sigma^2}{\mu} = \delta \mu + \frac{1 - \gamma}{\theta},$$

the TPL distribution is more scattered than the Tweedie distribution.

By further differentiation of (11), and re-parameterizing in μ and σ , it also follows that

$$m_3 = E[(X_{TPL} - \mu)^3] = \frac{2\sigma^4}{\mu} - \frac{(1 - \gamma)\sigma^2}{\theta} + \frac{(1 - \gamma)\mu}{\theta^2} \quad (24)$$

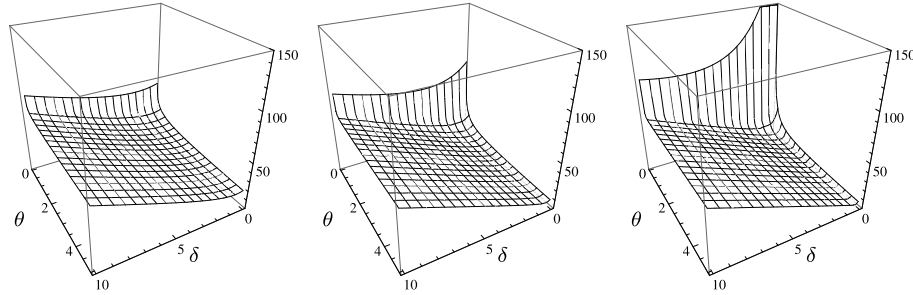


FIG 1. Plots of κ_4 as a function of (δ, θ) , in the case $\lambda = 1$ and for $\gamma = 0.25$ (left panel), $\gamma = 0.50$ (central panel) and $\gamma = 0.75$ (right panel).

and

$$m_4 = E[(X_{TPL} - \mu)^4] = \frac{6\sigma^6}{\mu^2} + \frac{(3\theta\mu - 6(1-\gamma))\sigma^4}{\theta\mu} + \frac{(5 - 6\gamma + \gamma^2)\sigma^2}{\theta^2} + \frac{(1-\gamma^2)\mu}{\theta^3}. \quad (25)$$

The moment expressions are useful because they provide further insight on the flexibility of the new distribution, e.g. through the skewness coefficient $\kappa_3 = m_3/\sigma^3$ and the kurtosis index $\kappa_4 = m_4/\sigma^4$. We sketch this flexibility in Figures 1 and 2, where we display several three-dimensional plots of κ_4 as a function of δ and θ , in the case $\lambda = 1$ and for a number of values of γ . It is clear from the plots that the new distribution substantially extends the range of the kurtosis index with respect to the Tweedie law. Indeed, for any given θ , κ_4 rapidly grows as δ increases for all the values of γ , either positive or negative. The reported feature is very important also for negative values of γ , since in this case a mixed distribution with a bold mass at zero and quite heavy tails can be achieved – in contrast with the Tweedie distribution. We argue that this property provides strong motivation for our proposal and shows its potential for application. Even if the new distribution involves four parameters, it displays a substantially extended morphology with respect to the Tweedie law – which is anyway a special case, being recovered in the limit as $\delta \rightarrow 0^+$. Empirical evidence of the gain obtained through the TPL distribution is provided in Section 6. We thus conclude that involving an extra parameter is worth the computational and statistical effort: our distribution is able to capture very marked deviations from normality, while keeping the original flexibility of the Tweedie model in terms of shape and modality. An additional probabilistic justification of our proposal within the framework of Lévy processes is given in the next section.

5. Connection with Lévy processes

Informally speaking, a non-negative Lévy process, say $X = (X_t)_{t \geq 0}$, is a stochastic jump process with non-negative, independent, time-homogeneous increments. The Laplace transform of such a process is given by the Lévy-Khintchine rep-

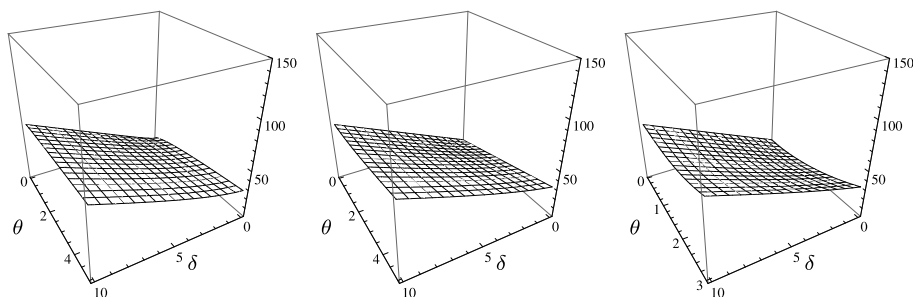


FIG 2. As Figure 1, but now for $\gamma = -0.5$ (left panel), $\gamma = -1$ (central panel) and $\gamma = -2$ (right panel).

resentation

$$L_{X_t}(s) = \exp(-t\psi_X(s)), \quad \text{Re}(s) > 0,$$

while the function

$$\psi_X(s) = \int_0^\infty (1 - e^{-sx})\nu_X(dx)$$

is the characteristic exponent of the Lévy process and ν_X is the Lévy measure (see Sato [41]). In what follows, we also consider subordination of the Lévy process X . Subordination is defined as a transformation of the original process X to a new stochastic process through random time change by a further non-negative Lévy process $Z = (Z_t)_{t \geq 0}$ independent of X and with Laplace transform

$$L_{Z_t}(s) = \exp(-t\psi_Z(s)), \quad \text{Re}(s) > 0,$$

where

$$\psi_Z(s) = \int_0^\infty (1 - e^{-sx})\nu_Z(dx)$$

is the characteristic exponent, while ν_Z is the Lévy measure. Hence, subordination produces a new non-negative Lévy process, say $Y = (Y_t)_{t \geq 0}$, where $Y_t = X_{Z_t}$, with Laplace transform given by

$$L_{Y_t}(s) = \exp[-t\psi_Z(\psi_X(s))], \quad \text{Re}(s) > 0.$$

We see subordination as an interesting modification of the root process for modelling real data, since many phenomena are likely to change more rapidly in different time segments. In such a case, it is preferable to represent the connection between calendar time and the pace of the phenomenon as random, so that a time-deformed process is achieved.

In order to provide further insight on the genesis of the TPL law, let us first assume that X is a TPS process with $\gamma \in]0, 1[$. In such a case, by considering the exponentially-tilted stable Lévy measure, i.e.

$$\frac{\nu_X(dx)}{dx} = \frac{\gamma\lambda}{\Gamma(1-\gamma)} x^{-\gamma-1} \exp(-\theta x), \quad x \in \mathbb{R}^+,$$

where $\lambda \in \mathbb{R}^+$ and $\theta \in \mathbb{R}^+$, we obtain the following characteristic exponent (Barabesi et al. [5])

$$\psi_X(s) = \lambda((\theta + s)^\gamma - \theta^\gamma).$$

Moreover, if Z is a Gamma process with Lévy measure

$$\frac{\nu_Z(dx)}{dx} = \eta x^{-1} \exp(-x/\delta), \quad x \in \mathbb{R}^+,$$

where $\delta \in \mathbb{R}^+$ and $\eta \in \mathbb{R}^+$, the corresponding characteristic exponent is given by

$$\psi_Z(s) = \eta \log(1 + \delta s).$$

Hence, subordination of X by Z gives rise to the Lévy process Y with Laplace transform

$$L_{Y_t}(s) = (1 + \lambda\delta((\theta + s)^\gamma - \theta^\gamma))^{-\eta t}, \quad \operatorname{Re}(s) > 0. \quad (26)$$

By letting $\eta t = 1/\delta$ for a fixed t , when $\gamma \in]0, 1[$ the TPL random variable proposed in our work may thus be considered as the “outcome” of subordination of a TPS process by a Gamma process.

On the other hand, when $\gamma \in \mathbb{R}^-$ the TPS process X may be seen as a compound Poisson process, in such a way that the primary Poisson process has rate ρ and each jump is assumed to be a Gamma random variable $X_G(-\gamma, 1/\theta)$. The corresponding Laplace transform is given by

$$L_{X_t}(s) = \exp[-\rho t(1 - (1 + s/\theta)^\gamma)], \quad \operatorname{Re}(s) > 0,$$

and X is a Lévy process with characteristic exponent

$$\psi_X(s) = \rho(1 - (1 + s/\theta)^\gamma) = \lambda(\theta^\gamma - (\theta + s)^\gamma),$$

where $\lambda = \rho/\theta^\gamma$. Thus, if Z is taken to be a Gamma process, subordination of X by Z gives rise to the Lévy process Y with Laplace transform

$$L_{Y_t}(s) = (1 - \lambda\delta((\theta + s)^\gamma - \theta^\gamma))^{-\eta t}, \quad \operatorname{Re}(s) > 0. \quad (27)$$

Therefore, by letting $\eta t = 1/\delta$ for a fixed t , when $\gamma \in \mathbb{R}^-$ the TPL random variable is again the “outcome” of a process arising from subordination of a TPS process – which is actually a compound Poisson process in this case – by a Gamma process.

In general, by suitably modifying the parametrization in the expression given by Vinogradov [44, Proposition 1.2], we write the Lévy measure of the TPS process for $\gamma \in]-\infty, 1[$ as

$$\frac{\nu_X(dx)}{dx} = \frac{|\gamma|\lambda}{\Gamma(1-\gamma)} x^{-\gamma-1} \exp(-\theta x), \quad x \in \mathbb{R}^+,$$

where $\lambda \in \mathbb{R}^+$ and $\theta \in \mathbb{R}^+$. In such a case, we readily obtain the following characteristic exponent

$$\psi_X(s) = \operatorname{sgn}(\gamma)\lambda((\theta + s)^\gamma - \theta^\gamma).$$

If Z denotes the Gamma process considered above, subordination of X by Z then yields the Lévy process Y with Laplace transform

$$L_{Y_t}(s) = (1 + \operatorname{sgn}(\gamma)\lambda\delta((\theta + s)^\gamma - \theta^\gamma))^{-\eta t}, \quad \operatorname{Re}(s) > 0.$$

Therefore, by letting $\eta t = 1/\delta$ for a fixed t , the TPL random variable is generally the “outcome” of subordination of a TPS process by a Gamma process.

We then derive the Lévy measure of the TPL process. When $\gamma \in]0, 1[$, on the basis of (19) and (26) by letting $\eta = 1/\delta$, the p.d.f. of the random variable Y_t may be written as

$$f_t(x) = \frac{\exp(-\theta x)}{x} \left(\frac{x^\gamma}{\lambda\delta}\right)^{t/\delta} \sum_{k=0}^{\infty} \frac{\binom{-t/\delta}{k}}{\Gamma(\gamma k + t\gamma/\delta)} \left(\frac{(1 - \lambda\delta\theta^\gamma)x^\gamma}{\lambda\delta}\right)^k, \quad x \in \mathbb{R}^+.$$

Hence, by using the method proposed by Barndorff-Nielsen [8] for computing Lévy densities, the Lévy measure of the TPL process may be given as

$$\begin{aligned} \frac{\nu_Y(dx)}{dx} &= \lim_{t \rightarrow 0^+} \frac{1}{t} f_t(x) = \frac{\gamma \exp(-\theta x)}{\delta x} \sum_{k=0}^{\infty} \frac{1}{\Gamma(\gamma k + 1)} \left(-\frac{(1 - \lambda\delta\theta^\gamma)x^\gamma}{\lambda\delta}\right)^k \\ &= \frac{\gamma \exp(-\theta x)}{\delta x} E_{\gamma,1} \left(-\frac{(1 - \lambda\delta\theta^\gamma)x^\gamma}{\lambda\delta}\right), \quad x \in \mathbb{R}^+. \end{aligned}$$

Moreover, when $\gamma \in \mathbb{R}^-$, on the basis of (21) and (27) by letting $\eta = 1/\delta$, the p.d.f. of the random variable Y_t conditional on the event $\{Y_t > 0\}$ may be also expressed as

$$h_t(x) = \frac{1}{(1 + \lambda\delta\theta^\gamma)^{t/\delta}} \frac{\exp(-\theta x)}{x} \sum_{k=1}^{\infty} \frac{\binom{-t/\delta}{k}}{\Gamma(-\gamma k)} \left(-\frac{\lambda\delta}{(1 + \lambda\delta\theta^\gamma)x^\gamma}\right)^k, \quad x \in \mathbb{R}^+.$$

By considering the remark by Applebaum [2, p.22], the same method thus provides the Lévy measure of the TPL process as

$$\begin{aligned} \frac{\nu_Y(dx)}{dx} &= \lim_{t \rightarrow 0^+} \frac{1}{t} h_t(x) = -\frac{\gamma \exp(-\theta x)}{\delta x} \sum_{k=1}^{\infty} \frac{1}{\Gamma(-\gamma k + 1)} \left(\frac{\lambda\delta}{(1 + \lambda\delta\theta^\gamma)x^\gamma}\right)^k \\ &= -\frac{\gamma \exp(-\theta x)}{\delta x} \left(E_{-\gamma,1} \left(\frac{\lambda\delta}{(1 + \lambda\delta\theta^\gamma)x^\gamma}\right) - 1\right), \quad x \in \mathbb{R}^+. \end{aligned}$$

As a final remark, we note that the TPL process Y can be easily simulated by considering its discrete approximation at times $t_0 < t_1 < \dots < t_n$. Indeed, in this case the following identity in distribution holds

$$Y_{t_i} - Y_{t_{i-1}} \stackrel{\mathcal{L}}{=} X_{Z_{t_i - t_{i-1}}},$$

where $i = 1, 2, \dots, n$. The increments of the discrete process can thus be generated as realizations of the random variables $X_{TPS}(\gamma, X_G(\lambda\delta, t_i - t_{i-1}), \theta)$. A

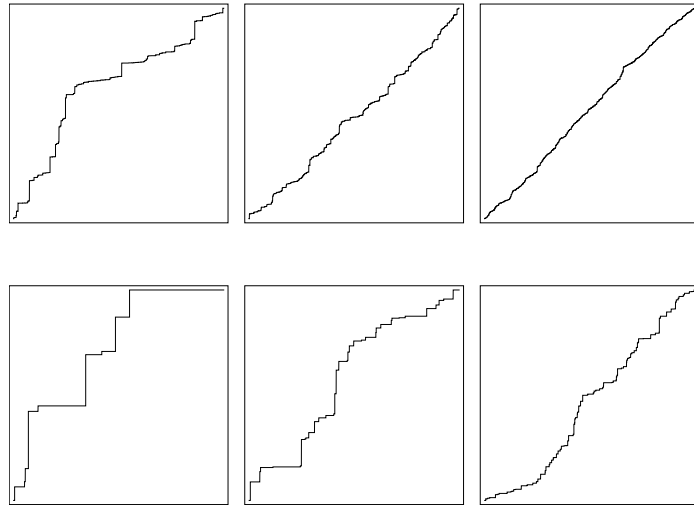


FIG 3. Randomly simulated paths of the proposed TPL process with 1000 steps and different parameter configurations. Top panels: $\gamma = 0.5$, $\lambda = 1.0$, $\theta = 1.0$ and $\delta = 0.3, 0.6, 2.0$. Bottom panels: $\gamma = -1.0$, $\lambda = 1.0$, $\theta = 1.0$ and $\delta = 0.5, 1.0, 2.0$.

few sample paths of Y with 1000 steps, obtained through this method, are displayed in Figure 3 for different parameter configurations. In particular, in the top panels we show three realizations where $\gamma = 0.5$, $\lambda = 1.0$, $\theta = 1.0$, and the shape parameter varies as $\delta = 0.3, 0.6, 2.0$. It is apparent that the realizations of the process are smoother as δ increases, since in this case the Lévy measure underlying the Gamma process tends to be more concentrated. On the other hand, the realizations displayed in the lower panels are for a negative value of γ , i.e. $\gamma = -1.0$, while $\lambda = 1.0$ and $\theta = 1.0$ as before, and $\delta = 0.5, 1.0, 2.0$. In this setting the realizations show trajectories with jumps, in accordance with the compound Poisson nature of the process. However, the jumps are more evident for small δ , while they reduce in size as δ increases, given the nature of the Gamma process subordination. Such pictures thus reinforce the idea that the proposed TPL process is considerably more flexible than the common TPS process and that it can become a valuable candidate for representing complex phenomena, even if it comes at the cost of an extra parameter. See also Rachev et al. [39, Ch.8] for more refined – albeit more cumbersome – techniques for random simulation of Lévy processes.

6. Application to international trade

We demonstrate the potential of our proposal by fitting the new TPL distribution to two data sets taken from international trade. Each data set contains the monthly aggregates of import quantities for a specific product that were registered in European Union (EU) Member States from non-EU countries in

the period August 2008 – July 2012. The information comes from the official extra-EU trade statistics, *Extrastat*, extracted from the COMEXT database of Eurostat. The monthly aggregates of imports are built from the customs declarations collected from individuals or companies by the Member States, following a strictly regulated process (Eurostat [17]). An overview of the resulting information archive is given in Barabesi et al. [5]. Sound statistical modeling of such data is an important tool when the task is to assess the performance of alternative anti-fraud methods; see, e.g., Barabesi et al. [4] and Cerioli and Perrotta [13] for a description of some of the statistical challenges that arise when investigating fraud patterns in international trade. The availability of flexible statistical models that adequately describe the distribution of trade quantities for a large number of products can also provide direct support to the EU policy makers, e.g., in the form of tools for monitoring the effect of policy measures.

Due to the combination of economic activities and normative constraints, the empirical distribution of traded quantities is often markedly skew with heavy tails, a large number of rounding errors in small-scale transactions due to data registration problems, and structural zeros arising because of confidentiality issues related to national regulations. Such features are typically not easy to analyze and Barabesi et al. [5] show that the Tweedie distribution leads to sensible models in several cases of interest. However, also the fit provided by the Tweedie distribution can be improved for some products, especially when a very heavy tail is present. This is the framework of our first example, concerning the product *Port (Wine of fresh grapes, including fortified wines, in containers holding 2 litres or less)*, within the Chapter of Beverages, Spirits and Vinegar. The empirical distribution of this product only contains positive values, but many of them are small and both skewness and kurtosis are impressively high. In particular, the empirical skewness coefficient is $\bar{\kappa}_3 = 5.9$ and the empirical kurtosis index is $\bar{\kappa}_4 = 47.9$, due to the presence of a small number of transactions involving more than 5000 litres (the maximum quantity is 14148 litres). To further appreciate the flexibility of our proposal, we then fit the TPL model to the traded quantities of *Fillets of Pangasius (Pangasius spp.), fresh or chilled* within the Chapter of Fish and Crustaceans, Molluscs and other Aquatic Invertebrates, a smaller data set for which both a mass of null values and a rather heavy tail ($\bar{\kappa}_4 = 14.3$) are present. Both examples provide relevant instances of products that may be affected by illegal trading behavior, such as counterfeiting, smuggling and misdeclaration of the product values or codes. Indeed, a major fraud due to systematic under-reporting of price for a fishery product is extensively described by Cerasa [12] and Riani et al. [40].

We perform computations through a Matlab algorithm that includes numerical evaluation of the Mittag-Leffler function and constrained optimization for Maximum Likelihood fitting. A good choice of the starting values of the parameters is important both for speeding up the algorithm and for reaching the best solution, as is often the case when maximizing complex multi-parameter functions. We have found that a sensible strategy is to combine random selection with moment matching. Specifically, starting from a randomly chosen value of γ and θ , say $\tilde{\gamma}_0$ and $\tilde{\theta}_0$, we first find the values $\tilde{\lambda}_0$ and $\tilde{\delta}_0$ that equate the theo-

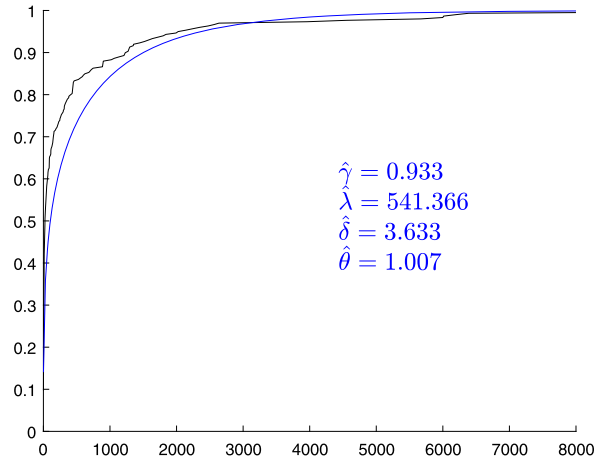


FIG 4. Empirical distribution function of product Port (Wine of fresh grapes, including fortified wines, in containers holding 2 litres or less) and fitted TPL distribution function.

retical mean (22) and variance (23) to the empirical moments. We then use the quadriplet $(\tilde{\gamma}_0, \tilde{\theta}_0, \tilde{\lambda}_0, \tilde{\delta}_0)$ as a starting point for minimizing the distance between the whole set of theoretical moments (22)–(25) and their empirical values. The resulting parameter estimates, say $(\tilde{\gamma}, \tilde{\theta}, \tilde{\lambda}, \tilde{\delta})$, are taken as the starting values of our Maximum Likelihood algorithm. An alternative, which is not considered in this paper, could be to obtain $(\tilde{\gamma}, \tilde{\theta}, \tilde{\lambda}, \tilde{\delta})$ by minimizing a crude and fast approximation of the L_2 -distance between the Laplace transform (11) and its empirical version.

In the *Port* data set we have $n = 299$ positive observations, with a mean of 453.622 litres and a standard deviation of 1350.63 litres. Therefore, $\gamma \in]0, 1]$ and, if (x_1, \dots, x_n) represents the observed sample, the likelihood function (with respect to the Lebesgue measure) is

$$\ell(\gamma, \lambda, \theta, \delta) = \prod_{i=1}^n \frac{\exp(-\theta x_i) x_i^{\gamma/\delta - 1}}{(\lambda \delta)^{1/\delta}} E_{\gamma, \gamma/\delta}^{1/\delta} \left(-\frac{(1 - \lambda \delta \theta^\gamma) x_i^\gamma}{\lambda \delta} \right).$$

Figure 4 shows the fit of the TPL model in our first example, by displaying the empirical distribution function of the data together with the estimate of the TPL distribution function obtained from (19) and restricted for clarity to values $x \in]0, 8000[$. This estimate is obtained through numerical integration. Our Matlab code is based on recursive calculation of the integral that stops when the error is $< 10^{-6}$. The Matlab version used for calculations is the R2015. The resulting parameter estimates $(\hat{\gamma}, \hat{\theta}, \hat{\lambda}, \hat{\delta})$ are also displayed in Figure 4. Comparison with the Tweedie distribution is provided in Table 1, which reports the log-likelihoods of the two competing models and gives the decrease in the criterion

$$\text{AIC} = -2 \log \hat{\ell} + 2\nu,$$

TABLE 1
Comparison of the fit provided by the TPL model and by the Tweedie distribution

	Log-likelihood TPL	Log-likelihood Tweedie (TPS)	Decrease in AIC index
Port	-1824.440	-2056.332	461.784
Fillets of Pangasius	-525.646	-530.039	6.786

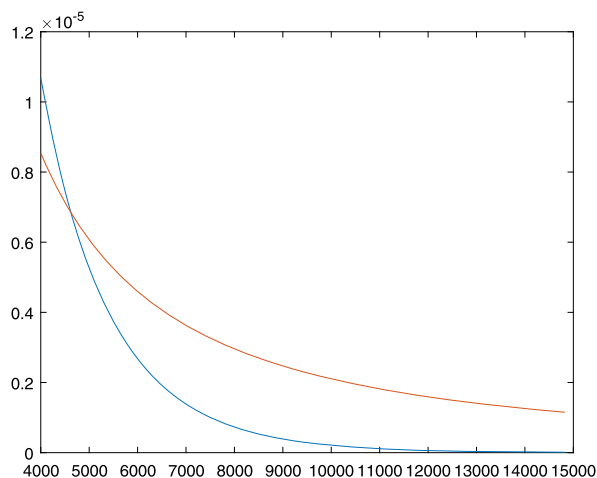


FIG 5. Zoom into the right-tail fit of the TPL (blue curve) and TPS (red curve) density functions to product Port (Wine of fresh grapes, including fortified wines, in containers holding 2 litres or less).

where $\hat{\ell}$ is the maximum of the likelihood function, when moving from the Tweedie distribution ($\nu = 3$) to our TPL proposal ($\nu = 4$). It is clearly seen that our four-parameter distribution yields a major decrease in both the likelihood function and the AIC index.

Further insight into the behavior of the two models under comparison can be obtained by looking at Figure 5, which shows a zoom into the right-hand tail of the estimated density functions for this product. The Maximum Likelihood estimates of the parameters in (6) are $\hat{\gamma} = 0.527$, $\hat{\lambda} = 9.529$ and $\hat{\theta} = 0.0001548$, using the algorithm described in Barabesi et al. [5]. We can see that the extremely large observed values of skewness and kurtosis lead the Tweedie distribution to become very close to a Positive Stable one, with virtually no tempering at all. It is thus apparent why our TPL model, which flexibly admits a very wide (but finite) range of values for (24) and (25), provides a preferable solution in this application. Indeed, the TPL model-based estimate of the kurtosis index results in $\hat{\kappa}_4 = 24.8$, which is reasonably close to the corresponding empirical coefficient, while we would obtain a value as high as $\hat{\kappa}_4 = 159.20$ under the Tweedie model. As anticipated by the panel of Figure 1 corresponding to the largest value of γ , when $\delta \rightarrow 0^+$ (i.e. in the limiting Tweedie case) high degrees of kurtosis can be reached only with $\theta = 0$. A much wider range of values for skewness and kurtosis measures is instead available through our model.

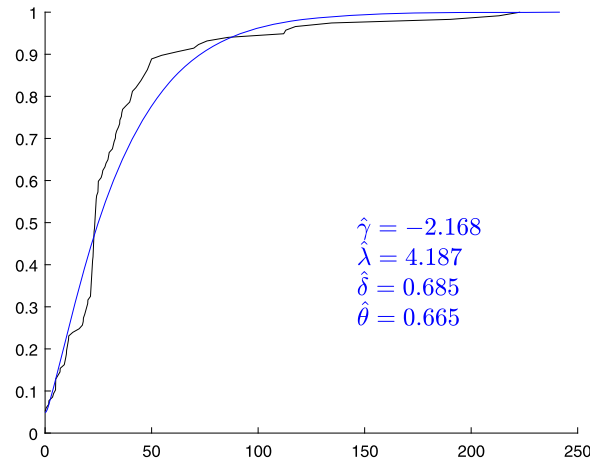


FIG 6. Empirical distribution function of product Fillets of Pangasius (*Pangasius* spp.), fresh or chilled and fitted TPL distribution function.

We obtain similar results also in our second example, for which we have $n = 117$ observations, $m = 6$ of which are exactly zero. The number of null observations is the realization of a Binomial random variable with n trials and success probability (20). We thus take $\gamma \in]-\infty, 0]$ and the likelihood function (with respect to the Dirac measure at zero and the Lebesgue measure) of the TPL model is

$$\begin{aligned} \ell(\gamma, \lambda, \theta, \delta) &= P(X_{TPL} = 0)^m (1 - P(X_{TPL} = 0))^{n-m} \prod_{i=1}^{n-m} h_{X_{TPL}}(x_i) \\ &= \varrho^{-n/\delta} \left(\frac{\lambda\delta}{\varrho} \right)^{n-m} \prod_{i=1}^{n-m} \frac{\exp(-\theta x_i)}{x_i^{\gamma+1}} \times \\ &\times \left\{ E_{-\gamma, -\gamma}^{1/\delta} \left(\frac{\lambda\delta}{\varrho x_i^\gamma} \right) + (\gamma - \gamma/\delta) E_{-\gamma, -\gamma+1}^{1/\delta} \left(\frac{\lambda\delta}{\varrho x_i^\gamma} \right) \right\}, \end{aligned}$$

where we have re-indexed the sample (x_1, \dots, x_n) in such a way that the first $(n - m)$ observations are non-null, while the remaining m observations are zero, and $\varrho = 1 + \lambda\delta\theta^\gamma$ as already defined in §4. Figure 6 shows the fit of our model, together with the corresponding parameter estimates. Although in this example the results obtained through the TPL model are qualitatively similar to those provided by the Tweedie distribution, they still show an improvement. Again, the AIC criterion reported in Table 1 conveys the finding that the addition of an extra parameter is more than compensated by the corresponding decrease in the likelihood function.

Acknowledgements

This work was partly supported by the 2016-2017 Institutional Program of the European Commission Joint Research Centre, and by the related Project “Automated Monitoring Tool on External Trade, Step5” of the European Anti-Fraud Office, under the Hercule-III Program. We warmly thank the Editor and one anonymous reviewer for very careful reading of our work and for several helpful comments that have considerably improved the presentation.

References

- [1] Aalen, O. (1992). Modelling heterogeneity in survival analysis by the compound poisson distribution. *Annals of Applied Probability* 2, 951–972. [MR1189425](#)
- [2] Applebaum, D. (2005). Lectures on Lévy processes, stochastic calculus and financial applications. 3me cycle de la Suisse Romande en Statistique et Probabilité Appliquée, Ovronnaz, Switzerland. <http://www.applebaum.staff.shef.ac.uk/talks.html>.
- [3] Barabesi, L., C. Becatti, and M. Marcheselli (2016). The tempered discrete Linnik distribution. *arXiv:1605.02326v1*.
- [4] Barabesi, L., A. Cerasa, A. Cerioli, and D. Perrotta (2016). Goodness-of-fit testing for the Newcomb-Benford law with application to the detection of customs fraud. *Journal of Business and Economic Statistics in press*, DOI:10.1080/07350015.2016.1172014. [MR3413561](#)
- [5] Barabesi, L., A. Cerasa, D. Perrotta, and A. Cerioli (2016). Modeling international trade data with the Tweedie distribution for anti-fraud and policy support. *European Journal of Operational Research* 248, 1031–1043. [MR3413561](#)
- [6] Barabesi, L. and L. Pratelli (2014). A note on a universal random variate generator for integer-valued random variables. *Statistics and Computing* 24, 589–596. [MR3223543](#)
- [7] Barabesi, L. and L. Pratelli (2015). Universal methods for generating random variables with a given characteristic function. *Journal of Statistical Computation and Simulation* 85, 1679–1691. [MR3315764](#)
- [8] Barndorff-Nielsen, O. (2000). Probability densities and Lévy densities. Technical report, Aarhus University, Centre for Mathematical Physics and Stochastics (MaPhySto). Research Report 18.
- [9] Barndorff-Nielsen, O. and N. Shephard (2002). Normal modified stable processes. *Theory of Probability and Mathematical Statistics* 65, 1–19. [MR1936123](#)
- [10] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Advances in Applied Probability* 31, 929–953. [MR1747450](#)
- [11] Carr, P., H. Geman, D. Madan, and M. Yor (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business* 75, 305–332.

- [12] Cerasa, A. (2016). Combining homogeneous groups of preclassified observations with application to international trade. *Statistica Neerlandica* 70, 229–259. [MR3520708](#)
- [13] Cerioli, A. and D. Perrotta (2014). Robust clustering around regression lines with high density regions. *Advances in Data Analysis and Classification* 8, 5–26. [MR3168677](#)
- [14] Christoph, G. and K. Schreiber (2001). Positive Linnik and discrete Linnik distributions. In N. Balakrishnan, I. Ibragimov, and V. Nevzorov (Eds.), *Asymptotic Methods in Probability and Statistics with Applications*, pp. 3–17. Boston: Birkhäuser. [MR1890314](#)
- [15] Cont, R. and P. Tankov (2004). *Financial Modeling With Jump Processes*. Boca Raton: Chapman and Hall. [MR2042661](#)
- [16] Devroye, L. and L. James (2014). On simulation and properties of the stable law. *Statistical Methods and Applications* 23, 307–343. [MR3233961](#)
- [17] Eurostat (2006). Statistics on the trading of goods – user guide. URL: <http://epp.eurostat.ec.europa.eu/>. ISSN 1725-0153, ISBN 92-79-01577-X.
- [18] Fattorini, L. and A. Lemmi (2006). The stochastic interpretation of the Dagum personal income distribution: a tale. *Statistica LXVI*, 325–329. [MR2414651](#)
- [19] Favaro, S., M. Lomeli, B. Nipoti, and Y. Teh (2014). On the stick-breaking representation of σ -stable Poisson-Kingman models. *Electronic Journal of Statistics* 8, 1063–1085. [MR3263112](#)
- [20] Favaro, S., B. Nipoti, and Y. Teh (2015). Random variate generation for Laguerre-type exponentially tilted α -stable distributions. *Electronic Journal of Statistics* 9, 1230–1242. [MR3355756](#)
- [21] Feller, W. (1971). *An Introduction to Probability Theory and its Applications. Vol. II. 2nd Edition*. New York: Wiley. [MR0270403](#)
- [22] Gorenflo, R., A. Kilbas, F. Mainardi, and S. Rogosin (2014). *Mittag-Leffler Functions, Related Topics and Applications*. Berlin: Springer. [MR3244285](#)
- [23] Grabchak, M. (2016). *Tempered Stable Distributions*. New York: Springer. [MR3468484](#)
- [24] Haubold, H., A. Mathai, and R. Saxena (2011). Mittag-Leffler functions and their applications. *Journal of Applied Mathematics* 2011, Article ID 298628, 51 pages. [MR2800586](#)
- [25] Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387–396. [MR0855898](#)
- [26] James, L. (2010). Lamperti-type laws. *Annals of Applied Probability* 20, 1303–1340. [MR2676940](#)
- [27] Johnson, N., A. Kemp, and S. Kotz (2005). *Univariate Discrete Distributions. 3rd Edition*. New York: Wiley. [MR2163227](#)
- [28] Jörgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society, Series B* 49, 127–162. [MR0905186](#)
- [29] Jose, K., P. Uma, V. Lekshmi, and H. Haubold (2010). Generalized Mittag-Leffler distributions and processes for applications in astrophysics and time series modeling. In H. Haubold and A. Mathai (Eds.), *Proceedings of the*

- Third UN/ESA/NASA Workshop on the International Heliophysical Year 2007 and Basic Space Science*. New York: Springer. [MR2723241](#)
- [30] Kanter, M. (1975). Stable densities under change of scale and total variation inequalities. *Annals of Probability* 3, 697–707. [MR0436265](#)
- [31] Klebanov, L. and L. Slámová (2015). Tempered distributions: does universal tempering procedure exist? *arXiv:1505.02068v1*.
- [32] Kleiber, C. (2008). A guide to the Dagum distributions. In D. Chotikapanich (Ed.), *Modeling Income Distributions and Lorenz Curves*, pp. 97–117. New York: Springer.
- [33] Lijoi, A. and I. Prunster (2014). Discussion of ‘On simulation and properties of the stable law’ by L. Devroye and L. James. *Statistical Methods and Applications* 23, 371–377. [MR3233966](#)
- [34] Lin, G. (2001). A note on the characterization of positive Linnik laws. *Australian & New Zealand Journal of Statistics* 43, 17–20. [MR1837492](#)
- [35] Linnik, Y. (1962). Linear forms and statistical criteria II. In *Translations in Mathematical Statistics and Probability* 3, pp. 41–90. Providence: American Mathematical Society. [MR0154361](#)
- [36] Mathai, A. and H. J. Haubold (2008). *Special Functions for Applied Scientists*. New York: Springer.
- [37] Pakes, A. (1995). Characterization of discrete laws via mixed sums and Markov branching processes. *Stochastic Processes and their Applications* 55, 285–300. [MR1313024](#)
- [38] Pillai, R. (1990). On Mittag-Leffler and related distributions. *Annals of the Institute of Statistical Mathematics* 42, 157–161. [MR1054728](#)
- [39] Rachev, S., Y. Kim, B. M. L., and F. Fabozzi (2011). *Financial Models with Lévy Processes and Volatility Clustering*. Hoboken, New Jersey: Wiley.
- [40] Riani, M., A. Cerioli, A. Atkinson, D. Perrotta, and F. Torti (2008). Fitting robust mixtures of regression lines with the forward search. In *Mining Massive Datasets for Security Applications*, pp. 271–286. IOS Press.
- [41] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge, UK: Cambridge University Press.
- [42] Steutel, F. and K. van Harn (2004). *Infinite Divisibility of Probability Distributions on the Real Line*. New York: Dekker. [MR2011862](#)
- [43] Tweedie, M. (1984). An index which distinguishes between some important exponential families. In J. Ghosh and J. Roy (Eds.), *Statistics: Applications and New Directions, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, pp. 579–604. Calcutta: Indian Statistical Institute. [MR0786162](#)
- [44] Vinogradov, V. (2004). On the power-variance family of probability distributions. *Communications in Statistics – Theory and Methods* 33, 1007–1029. [MR2054855](#)
- [45] Zolotarev, V. (1986). *One-dimensional Stable Distributions*. Providence: Translations of Mathematical Monographs 65, American Mathematical Society. [MR0854867](#)