WESTMINSTER
BUSINESS SCHOOL

WORKING PAPER SERIES

2021/001

Polynomial adjusted Student-t densities for modeling asset returns

Ángel León
Dept. de Fundamentos del Análisis Económico (FAE),
Universidad de Alicante

Trino-Manuel Ñíguez
School of Organisations, Economy and Society,
Westminster Business School, University of Westminster

All rights reserved

UNIVERSITY OF
VISION
STRATEGY
OPPORTUNITY
WESTMINSTER

Polynomial adjusted Student-t densities for modeling asset returns

Ángel León
Dept. Fundamentos del Análisis Económico (FAE), Universidad de Alicante, Alicante 03690, Spain.
E-mail: aleon@ua.es

Trino-Manuel Ñíguez*
Westminster Business School, University of Westminster, London NW1 5UL, UK.
E-mail: t.m.niguez@wmin.ac.uk

Abstract

We present a polynomial expansion of the standardized Student-t distribution. Our density, obtained through the polynomial adjusted method in Bagnato, Potì and Zoia (2015), is an extension of the Gram-Charlier density in Jondeau and Rockinger (2001). We derive the closed-form expressions of the moments, the distribution function and the skewness-kurtosis frontier for a well-defined density. An empirical application is also implemented for modeling heavy-tailed and skewed distributions for daily asset returns. Both in-sample and backtesting analysis show that this new density can be a good candidate for risk management.

Keywords: Backtesting; Expected shortfall; Kurtosis; Orthogonal polynomials; Skewness, VaR

JEL classification codes: C1, C22, G11.

*Corresponding author

1 Introduction

The Gram-Charlier (GC) distribution studied, among others, in Corrado and Su (1996), Jondeau and Rockinger (2001, henceforth JR), Del Brio and Perote (2012), Schlögl (2013), León and Moreno (2017) and Zoia, Biffi and Nicolussi (2018) has become popular in financial economics as a generalization of the normal distribution. The GC density is the expansion of the standard normal density based on orthogonal polynomials.[1] It can be seen as a particular case of the polynomially adjusted (PA) class of densities of Bagnato, Potì and Zoia (2015, henceforth BPZ) and so, the GC is also referred to as the PA Gaussian (PAG). These authors show a simple theorem that links the higher-order moments (skewness and kurtosis) of a polynomially expanded (parent) distribution to those of the target distribution. The selected parent density tends to verify the properties of symmetry and unimodality and as a result, the associated PA density presents advantages in terms of theoretical tractability and empirical applicability. The two parameters of the polynomial expansion become directly the skewness and excess kurtosis implied in the PA density, indeed. The orthogonality of the polynomials depends on the selected parent density. For instance, the Hermite polynomials verify the orthogonality condition when the parent density is the Gaussian. BPZ (2015) also consider, as parent densities, the standardized forms of both the hyperbolic secant and logistic densities to obtain their corresponding PA densities referred to as PAHS and PAL, respectively.[2]

We present the polynomial expansion of the standardized Student-t (T) distribution, or PA Student-t (PAST hereafter). The PAST nests the PAG as the T distribution's degrees of freedom tend to infinity and hence, it becomes more flexible than the latter for modeling the heavy-tailed and skewed distributions of asset returns. It is well known that the PAG density is only suitable for return series with moderate kurtosis. The literature on this topic has developed very quickly during recent years, see for instance, Wilhelmsson (2006); Komunjer (2007); Bali, Mo and Tang (2008); Zhu and Galbraith (2011); Dendramis, Spungin and Tzavalis (2014); Harvey and Sucarrat (2014); Li, Liu and Ng (2015); Feunou, Jahan-Parvar and Tédongap (2016); Kumar and Patil (2016); León and Ñiguez (2020) and Thiele (2020). Contributing to this line of research, we show that the PAST can be a good candidate for statistical analysis of financial returns and, in general, for fitting series with high levels of kurtosis. Another advantage of using PAST is that it nests the standard normal (N) which is an advantage in terms of tractableness.

We illustrate the practical use of the PAST pdf through an application to modeling the conditional distribution of asset returns. For that purpose, we implement the popular conditional variance model by Glosten-Jagannathan-Runkle (1993, henceforth GJR). We test the performance of our model through in-sample and out-of-sample (OOS) analyses. The data we use comprises three stock indexes, four exchange rates, and three commodity indexes. For comparison purposes, we consider the following alternative densities: both N and T as benchmark densities; PAL, PAHS and GC as alternative PA densities, and the skewed-t (sk-

[1] See Chihara (1978) for an introduction to orthogonal polynomials.

[2] These PA densities are also known as Gram-Charlier-like (GC-like) expansions. See Nicolussi and Zoia (2020) for the case of multivariate GC-like expansions.

T) density of Hansen (1994). Models relative forecasting performance is evaluated through the backtesting methods for Value-at-Risk (VaR) and expected shortfall (ES) of Du and Escanciano (2017).

The remainder of the paper is structured as follows. Section 2 deals with the general framework to construct PA densities according to the results in BPZ (2015) and provides some new statistical properties. In Section 3 we present the PAST distribution and specifically, we obtain the probability density function (pdf), cumulative distribution function (cdf) and study the parametric properties together with its skewness-kurtosis frontier that guarantees the non-negativeness of the PAST density. Section 4 provides an empirical application to asset returns and a performance comparative analysis. Conclusions are given in Section 5. The proofs for some results in Section 3 are collected in the Appendix.

2 Polynomial adjusted distributions

2.1 Density function

Let $f(x, \boldsymbol{\delta})$ denote a parent standardized symmetric and unimodal density function such that $x \in \mathbb{R}$ and $\boldsymbol{\delta}$ is the parameter vector implied in the density, then BPZ (2015) shows that its corresponding PA density is obtained as

$$g(x; \boldsymbol{\delta}, \boldsymbol{\theta}) = f(x; \boldsymbol{\delta})\, \psi(x; \boldsymbol{\delta}, \boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\theta} = (\theta_3, \theta_4)$ is another parameter vector, and $\psi(\cdot)$ is defined as

$$\psi(x; \boldsymbol{\delta}, \boldsymbol{\theta}) = 1 + \frac{\theta_3}{\gamma_3} p_3(x) + \frac{\theta_4}{\gamma_4} p_4(x), \tag{2}$$

such that $p_j(x)$ is a polynomial of order j with

$$p_3(x) = x^3 - a_1 x, \quad p_4(x) = x^4 - a_2 x^2 + a_3, \tag{3}$$

and

$$a_1 = m_4, \quad a_2 = \frac{m_6 - m_4}{m_4 - 1}, \quad a_3 = \frac{m_6 - m_4^2}{m_4 - 1}, \tag{4}$$

where $m_k = E_f\left[x^k\right]$ denotes the k-th non-central moment of x with $f(x, \boldsymbol{\delta})$ as pdf. Note that $m_k = m_k(\boldsymbol{\delta})$, then $a_i = a_i(\boldsymbol{\delta})$. Both polynomials in (3) with coefficients in (4) verify the orthogonality condition, i.e. $E_f[p_3(x)\, p_4(x)] = 0$, and $E_f[p_j(x)] = 0$. Finally, the coefficient $\gamma_j = \gamma_j(\boldsymbol{\delta})$ is the squared norm associated to $p_j(x)$, i.e. $\gamma_j = E_f\left[p_j^2(x)\right] = E_f\left[x^j p_j(x)\right]$, then

$$\gamma_3(\boldsymbol{\delta}) = m_6 - a_1 m_4, \quad \gamma_4(\boldsymbol{\delta}) = m_8 - a_2 m_6 + a_3 m_4. \tag{5}$$

2.2 Non-central moments

The selected parent density $f(x; \boldsymbol{\delta})$ must verify that its moments of all order exist or, at least, those necessary to compute a_i in (4). Since $m_1 = 0$, $m_2 = 1$ and $m_k = 0$ for k odd due to the symmetry of $f(\cdot)$, then

the first non-central moments from (1), denoted as $\widetilde{m}_k = E_g\left[x^k\right]$, are given by $\widetilde{m}_1 = 0$, $\widetilde{m}_2 = 1$, $\widetilde{m}_3 = \theta_3$ and $\widetilde{m}_4 = m_4 + \theta_4$. Thus, θ_3 and θ_4 become directly the skewness and excess kurtosis of $g(x;\boldsymbol{\delta},\boldsymbol{\theta})$, and $\psi(x;\boldsymbol{\delta},\boldsymbol{\theta})$ is interpreted as a polynomial shape-adapter function reshaping $f(x;\boldsymbol{\delta})$ for skewness via θ_3 and excess kurtosis via θ_4. Note that $E_f\left[x^k p_4(x)\right] = 0$ if k is odd, while $E_f\left[x^k p_3(x)\right] = 0$ if k is even, then the general expression of $\widetilde{m}_k$ for $k \in \mathbb{N}$ is given by

$$\widetilde{m}_k(\boldsymbol{\delta},\boldsymbol{\theta}) = \begin{cases} m_k(\boldsymbol{\delta}) + \theta_4\gamma_4^{-1}E_f\left[x^k p_4(x)\right], & k \text{ even} \\ \theta_3\gamma_3^{-1}E_f\left[x^k p_3(x)\right], & k \text{ odd.} \end{cases} \tag{6}$$

In short, $\widetilde{m}_k$ only depends on the skewness θ_3 (excess kurtosis θ_4) when k is odd (even).

2.3 Cumulative distribution function and Expected shortfall

Consider the cumulative distribution function or cdf $G(x;\boldsymbol{\delta},\boldsymbol{\theta}) = \int_{-\infty}^{x} g(u;\boldsymbol{\delta},\boldsymbol{\theta})\,du$, then

$$G(x;\boldsymbol{\delta},\boldsymbol{\theta}) = \sum\nolimits_{j=0}^{4}\omega_j\xi_j(x;\boldsymbol{\delta}) = \sum\nolimits_{j=0}^{4}\omega_j\int_{-\infty}^{x}u^j f(u;\boldsymbol{\delta})\,du, \tag{7}$$

where

$$\omega_0 = 1 + \theta_4\gamma_4^{-1}a_3, \quad \omega_1 = -\theta_3\gamma_3^{-1}a_1, \quad \omega_2 = -\theta_4\gamma_4^{-1}a_2, \quad \omega_3 = \theta_3\gamma_3^{-1}, \quad \omega_4 = \theta_4\gamma_4^{-1}. \tag{8}$$

Note that $\xi_0(x;\boldsymbol{\delta}) = F(x;\boldsymbol{\delta})$ is the cdf of $f(x;\boldsymbol{\delta})$, and $\xi_j(x;\boldsymbol{\delta})$ for $j \geq 1$ can be expressed in terms of one-sided truncated j-th moments, i.e. $\xi_j(x;\boldsymbol{\delta}) = F(x;\boldsymbol{\delta})\,E_{f_\star}\left[u^j\,|u \leq x\right]$ with $f_\star(u;\boldsymbol{\delta}) = f(u;\boldsymbol{\delta})/F(x;\boldsymbol{\delta})$ as pdf.

The expected shortfall, or ES, corresponding to the random variable x with $g(\cdot)$ as the pdf in (1) is easily obtained as

$$E_g\left[x\,|x \leq x_\alpha\right] = \frac{1}{\alpha}\int_{-\infty}^{x_\alpha} x g(x;\boldsymbol{\delta},\boldsymbol{\theta})\,dx = \sum\nolimits_{i=1}^{5}\varpi_i\xi_i(x_\alpha;\boldsymbol{\delta}), \tag{9}$$

such that $x_\alpha = G^{-1}(\alpha)$ is the α-quantile, or VaR, i.e. $G^{-1}(\alpha) = \inf\{x\,|G(x,\boldsymbol{\theta}) \geq \alpha\}$ with $G(\cdot)$ as the cdf in (7), and

$$\varpi_1 = \frac{1}{\alpha}\left(1 + \frac{\theta_4 a_3}{\gamma_4}\right), \quad \varpi_2 = -\frac{\theta_3 a_1}{\alpha\gamma_3}, \quad \varpi_3 = -\frac{\theta_4 a_2}{\alpha\gamma_4}, \quad \varpi_4 = \frac{\theta_3}{\alpha\gamma_3}, \quad \varpi_5 = \frac{\theta_4}{\alpha\gamma_4}. \tag{10}$$

Let $y = \mu + \sigma x$ be a linear transformation of x where $\mu \in \mathbb{R}$ and $\sigma > 0$ denote, respectively, the location and scale parameters. It is verified that $E\left[y\,|y \leq y_\alpha\right] = \mu + \sigma E_g\left[x\,|x \leq x_\alpha\right]$ with $y_\alpha = \mu + \sigma x_\alpha$.

2.4 Positivity restrictions

For a correct PA density in (1), we need to ensure the positivity of $\psi(x;\boldsymbol{\delta},\boldsymbol{\theta})$. There are two different methods. First, the one suggested by Gallant and Nychka (1987) consisting on squaring $\psi(x;\boldsymbol{\delta},\boldsymbol{\theta})$.[3] However, by doing so, one loses the interpretation of the parameters θ_j in (2), as the skewness and excess kurtosis, since we would face a new pdf: $h(x;\boldsymbol{\delta},\boldsymbol{\theta}) = \lambda f(x;\boldsymbol{\delta})\,\psi^2(x;\boldsymbol{\delta},\boldsymbol{\theta})$ with $\lambda^{-1} = \int_{-\infty}^{+\infty} h(x;\boldsymbol{\delta},\boldsymbol{\theta})\,dx = 1 + \theta_3^2\gamma_3^{-1} + \theta_4^2\gamma_4^{-1}$.

[3] See León, Mencía and Sentana (2009) for the parametric properties of the Gallant-Nychka density.

Second, the method suggested in JR (2001) for building the domain, or restricted parameter set, $\mathbf{\Lambda}$, over which $\psi(x;\boldsymbol{\delta},\boldsymbol{\theta}) > 0$ for every x. In this study, we follow the latter approach and leave the former for further research. So, we are interested in the boundary of $\mathbf{\Lambda}$ by using the analytical geometry concept of the *envelope*. For a given x, the skewness-kurtosis frontier (SKF) that guarantees positivity must satisfy the following equation system: $\psi(x;\boldsymbol{\delta},\boldsymbol{\theta}) = 0$ and $\partial\psi/\partial x = 0$. As a result, we obtain the SKF expressions for θ_3 and θ_4 as functions of x, which are given by

$$\theta_3(x;\boldsymbol{\delta}) = -\gamma_3 \frac{4x^3 - 2a_2 x}{x^6 + \beta_4 x^4 + \beta_2 x^2 + \beta_0}, \quad \theta_4(x;\boldsymbol{\delta}) = \gamma_4 \frac{3x^2 - a_1}{x^6 + \beta_4 x^4 + \beta_2 x^2 + \beta_0}, \tag{11}$$

where $\beta_4 = a_2 - 3a_1$, $\beta_2 = a_1 a_2 - 3a_3$ and $\beta_0 = a_1 a_3$. In short, the PA density in (1) will be subject to $\boldsymbol{\theta} \in \mathbf{\Lambda}(\boldsymbol{\delta})$ such that any point in the frontier of $\mathbf{\Lambda}(\boldsymbol{\delta})$ is obtained according to (11). In the following section, we use $\mathbf{\Lambda}(\boldsymbol{\delta}) = \mathbf{\Lambda}(v)$ where v denotes the degrees of freedom for the standardized Student-t as the parent density.

3 The PAST distribution

3.1 Density function

The standardized Student-t density, or T density, is given by

$$f(x;v) = \frac{1}{(v-2)^{1/2} B(v/2, 1/2)} \left(1 + \frac{x^2}{v-2}\right)^{-(v+1)/2} \tag{12}$$

where $B(a,b) = \int_0^1 u^{a-1}(1-u)^{b-1} du$ with $a, b > 0$ is the ordinary beta function. The non-central moments only exist for $0 < j < v$ verifying that $m_1 = 0$, $m_2 = 1$, $m_{2k+1} = 0$ for $j = 2k+1$, and

$$m_{2k}(v) = (v-2)^k \prod_{i=1}^{k} \left(\frac{2i-1}{v-2i}\right) \tag{13}$$

for $j = 2k$. The PAST pdf is given by (1), then $g(x;v,\boldsymbol{\theta}) = f(x;v)\,\psi(x;v,\boldsymbol{\theta})$ with $f(\cdot)$ as the parent density in (12) and $\psi(\cdot)$ is the polynomial function in (2) subject to $v > 8$ so as to obtain $\gamma_j(v)$ in (5). For instance, if we take $v = 10$, then a_i in (4) and γ_j are given by $a_1 = 4$, $a_2 = 12$, $a_3 = 8$, $\gamma_3 = 24$ and $\gamma_4 = 672$.

For $v \to \infty$, we obtain the GC pdf: $g_{GC}(x;\boldsymbol{\theta}) = \phi(x)\,\psi_{GC}(x;\boldsymbol{\theta})$ where $\phi(x)$ is the standard normal pdf and $\psi_{GC}(x;\boldsymbol{\theta}) = 1 + \frac{\theta_3}{3!}H_3(x) + \frac{\theta_4}{4!}H_4(x)$ with $H_j(x)$ as the j-th (orthogonal) Hermite polynomial. The first four Hermite polynomials are: $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x$ and $H_4(x) = x^4 - 6x^2 + 3$.

Figure 1 provides an illustration of the PAST density with $v = 15$ (PAST15) as it compares to the sk-T and PAHS, plots of zoomed lower tails are also provided. All these densities have zero mean, unit variance and the same levels of skewness (-0.47) and kurtosis (11). More details about these latter values and the densities' parameters can be seen below in section 3.3.

[Figure 1]

3.2 Cdf and ES

The PAST cdf is given by $G(x;v,\boldsymbol{\theta})$ in (7) where $\xi_j(x;v) = \int_{-\infty}^{x} u^j f(u;v)\,du$ with $f(\cdot)$ as the pdf in (12). Next, we obtain the corresponding expressions of $\xi_j(x;v)$.

Corollary 1. *Let $\xi_0(x;v) = F(x;v)$ be the cdf of the standardized ST distribution, then*

$$F(x;v) = 1/2\,\left(1 + sgn(x)\,I_{\eta(x)}(1/2, v/2)\right), \tag{14}$$

where $sgn(x) = x/|x|$ for $x \neq 0$ (and 0, for $x = 0$), $\eta(x) = x^2/(x^2+v-2)$, $I_x(a,b) = B(x;a,b)/B(a,b)$ is the regularized beta function, and $B(x;a,b) = \int_0^x u^{a-1}(1-u)^{b-1}\,du$ is the incomplete beta function.

Proof. See Appendix. ■

Proposition 1. *The closed-form expression of $\xi_j(x;v)$ when j is an even number is obtained as*

$$\xi_{2k}(x;v) = \frac{m_{2k}(v)}{2} + sgn(x)\,(v-2)^k\,\frac{B\left(\frac{v}{2}-k, k+\frac{1}{2}\right)}{2B\left(\frac{v}{2},\frac{1}{2}\right)} I_{\eta(x)}\left(k+\frac{1}{2}, \frac{v}{2}-k\right), \tag{15}$$

with $m_{2k}(v)$ in (13).

Proof. See Appendix. ■

Proposition 2. *The closed-form expression of $\xi_j(x;v)$ when j is an odd number is obtained as*

$$\xi_{2k+1}(x;v) = (v-2)^{k+\frac{1}{2}}\,\frac{B\left(\frac{v-1}{2}-k, k+1\right)}{2B\left(\frac{v}{2},\frac{1}{2}\right)}\left[I_{\eta(x)}\left(k+1, \frac{v-1}{2}-k\right) - 1\right]. \tag{16}$$

Proof. See Appendix. ■

Hence, the cdf $G(x;v,\boldsymbol{\theta})$ in (7) for PAST (henceforth, $G_{PAST}(x;v,\boldsymbol{\theta})$) is computed given the equations (14), (15) and (16), and the expressions of ω_j in (8). For $v \to \infty$, we obtain the GC cdf: $G_{GC}(x;\boldsymbol{\theta}) = \Phi(x) - \frac{\theta_3}{3!}H_2(x)\,\phi(x) - \frac{\theta_4}{4!}H_3(x)\,\phi(x)$ where $\Phi(x)$ is the standard normal cdf.[4]

Figure 2 showcases the cdf and ES of PAST when $v = 15$ (PAST15) and $v = 300$ (PAST300), and both with zero mean, unit variance, skewness $\theta_3 = -0.5$ and excess kurtosis $\theta_4 = 2.4545$. The kurtosis of the parent T pdf, according to (13) for $k = 2$, is 3.5455 for the case of $v = 15$ and 3.0121 for $v = 300$. Note that the latter case converges to the parent N pdf and hence, to the GC pdf with the same above values of θ_3 and θ_4. In short, the skewness is the same for both PAST pdfs, and the kurtosis is 6 for PAST15 and 5.46 for PAST300. The plotted ES values for both PAST pdfs are computed according to the equations (9), (15), (16) and the expressions of ϖ_i in (10). The ES values are displayed for the range -4 to -3 where $G_{PAST}(-3;15,\boldsymbol{\theta}) = 0.0096$, $G_{PAST}(-3;300,\boldsymbol{\theta}) = 0.0125$, $G_{PAST}(-4;15,\boldsymbol{\theta}) = 0.0029$, and $G_{PAST}(-4;300,\boldsymbol{\theta}) = 0.001$ with $\boldsymbol{\theta} = (-0.5, 2.4545)$. We find that the ES (dashed) line for PAST300 is above the PAST15 one, which means that expected losses (i.e., ES with a negative sign) are greater the lower is

[4] See Proposition 2 in León and Moreno (2017) where the GC cdf is obtained as the definite integral of the GC pdf over $(-\infty, x]$. Note that the Hermite polynomials are expressed as orthonormal, instead of orthogonal, ones.

the parameter v. So, higher kurtosis levels for the same skewness levels lead to higher expected losses when we move deep inside the left tail.

[Figure 2]

3.3 Skewness-kurtosis frontier

Figure 3 exhibits several frontiers of $\mathbf{\Lambda}(v)$ by using both equations in (11) for the PAST pdf with $v \in \{9, 10, 15, 20\}$ and $v \to \infty$ for the PA Gaussian, or GC, pdf. More precisely, the frontiers are in terms of kurtosis instead of excess kurtosis. In short, the points of the skewness-kurtosis frontier (SKF) are given by $(sk, ku) = (\theta_3, m_4(v) + \theta_4)$ with m_4 in (13). Note that the SKFs are symmetric with respect to the x-axis. It is shown that the lower v, the higher ku and the lower sk, so the PAST's SKF enlarges significantly the GC's. For instance, any point from the frontier of $\mathbf{\Lambda}(\infty)$, or GC frontier, verifies that $3 \leq ku \leq 7$, $|sk| \leq 1.0493$, and the maximum size for sk is reached for $\theta_4 = 2.4508$. The equations in (11) for $v \to \infty$ are: $\theta_3(x) = -24H_3(x)/q(x)$ and $\theta_4(x) = 72H_2(x)/q(x)$ where $q(x) = x^6 - 3x^4 + 9x^2 + 9$. See also JR (2001).

[Figure 3]

In order to illustrate the effect that the PAST envelope for different ν has on the shape of the distribution, Figure 4 plots theoretical quantiles from zero mean and unit variance densities of the sk-T against PAL, PAHS and PAST with fixed $\nu = 10, 15$ (PAST10 and PAST 15, respectively). We showcase the comparison for both distribution tails: lower tail (quantiles from 0.001 to 0.05) and upper tail (quantiles from 0.95 to 0.999). Specifically, we employ a total of fifty equally-spaced quantiles in each tail. The PA distributions' parameters are fine-tuned so that their values of sk and ku match those of the sk-T. The specific higher-order moment values are $sk_0 = -0.4672$ and $ku_0 = 10.9588$ (rather similar to those empirical values exhibited below in Table 1; see also Figure 1 for an illustration of these densities).[5] This means that $\theta_3 = -0.4672$ for the three PA distributions, whilst $\theta_4 = 6.7588$ for PAL (i.e., $m_4 = 4.2$ for the parent logistic pdf) and $\theta_4 = 5.9588$ for PAHS (i.e., $m_4 = 5$ for the parent hyperbolic secant pdf). Note that the PAST, contrary to PAL and PAHS, can provide the level of ku_0 under different combinations of v and θ_4 as exhibited in Figure 3. We consider, for example, only two cases, namely, $\theta_4 = 6.9588$ for PAST with $v = 10$ (i.e., $m_4 = 4$ for the parent T pdf with $v = 10$), and $\theta_4 = 7.4133$ for PAST with $v = 15$ (i.e., $m_4 = 3.5455$ for the parent T with $v = 15$). Note that PAST10 becomes more flexible to capture higher kurtosis levels than PAST15, when both are restricted to the same skewness level of -0.47, as displayed in Figure 3. Several conclusions are suggested from Figure 4. First, we find that PAST10 provides much closer quantiles to those of the sk-T than PAST15 for both tails. That is, PAST10 quantiles are closer to the 45-degree dashed line. Second, PAL and PAHS quantiles are very similar for the lower tail, whilst PAHS quantiles are nearer to sk-T for the upper tail. Third, it is verified that more significant differences among PA quantiles (respecting sk-T

[5] Both sk_0 and ku_0 are obtained by plugging the sk-T parameters of $\lambda = -0.1$ (parameter controlling skewness) and $v = 4.8$ (degrees of freedom) into the skewness and kurtosis equations of the sk-T pdf in JR (2003).

ones) can be seen at the end of both tails. Finally, PAL and PAST15 seem to provide similar quantiles for both tails.

[Figure 4]

3.4 Conditional distribution of asset returns

Let r_t be the asset return process characterized by the sequence of conditional densities $h(r_t|I_{t-1};\boldsymbol{\Upsilon})$, where I_{t-1} denotes the information set available prior to the realization of r_t, $\boldsymbol{\Upsilon}=(\mu,\boldsymbol{\varsigma},v,\boldsymbol{\theta})$ is the vector of unknown parameters with μ as the constant mean of r_t, $\boldsymbol{\varsigma}$ is the subset characterizing the conditional variance of r_t, and $(v,\boldsymbol{\theta})$ characterize the shape of the PAST distribution innovations, z_t. Thus, the asset return model is $r_t=\mu+\varepsilon_t$ with $\varepsilon_t=\sigma_t z_t$ such that $\sigma_t^2=E\left[\varepsilon_t^2|I_{t-1}\right]$ is the GJR model:

$$\sigma_t^2=\alpha_0+\beta\sigma_{t-1}^2+\alpha_1^+\left(\varepsilon_{t-1}^+\right)^2+\alpha_1^-\left(\varepsilon_{t-1}^-\right)^2, \tag{17}$$

such that $\alpha_0>0$, $\beta\geq 0$, $\alpha_1^+\geq 0$, $\alpha_1^-\geq 0$, and consider $\varepsilon_t^+=\max(\varepsilon_t,0)$, $\varepsilon_t^-=\min(\varepsilon_t,0)$. Henceforth, the above process for r_t is referred to as GJR-PAST model. Note that the GJR nests the GARCH when $\alpha_1^+=\alpha_1^-$. Following He and Terasvirta (1999), we can rewrite the GJR as $\sigma_t^2=\alpha_0+c_t\sigma_{t-1}^2$ with $c_t=\beta+\alpha_1^+\left(z_{t-1}^+\right)^2+\alpha_1^-\left(z_{t-1}^-\right)^2$. If we assume (17) to be covariance stationary, then the unconditional variance of ε_t is $E\left(\varepsilon_t^2\right)=E\left(\sigma_t^2\right)=\alpha_0\left(1-E\left(c_t\right)\right)^{-1}$ such that $E\left(c_t\right)=\beta+\left(\alpha_1^-+\alpha_1^+\right)/2<1$. Since $E\left[r_t|I_{t-1}\right]=\mu$, then $E\left(\varepsilon_t^2\right)$ is the unconditional variance of r_t. Let $l_t=\ln h(r_t|I_{t-1};\boldsymbol{\Upsilon})$ be the log-likelihood (LL) function for a particular observation r_t, then

$$l_t=-\frac{1}{2}\ln\left(\sigma_t^2\right)-\frac{1}{2}\ln(v-2)-\ln B\left(\frac{v}{2},\frac{1}{2}\right)-\frac{v+1}{2}\ln\left(1+\frac{z_t^2}{v-2}\right)+\ln\psi(z_t;v,\boldsymbol{\theta}), \tag{18}$$

where $z_t=(r_t-\mu)/\sigma_t$. Finally, alternative distributions for z_t are also implemented here for a robustness analysis. Thus, $z_t\sim i.i.d.\ D(0,1)$ such that $D(0,1)$ denotes a specific density (N, T, sk-T, GC, PAST, PAHS and PAL) with zero mean and unit variance, and l_t^D as the corresponding LL per observation. Specifically, the expression of l_t^D with a general PA density (henceforth, l_t^{PA}) under both constant mean and GJR structure in (17) is easily obtained as

$$l_t^{PA}=-\frac{1}{2}\ln\left(\sigma_t^2\right)+\ln f(z_t;\boldsymbol{\delta})+\ln\psi(z_t;v,\boldsymbol{\theta}), \tag{19}$$

with $z_t=(r_t-\mu)/\sigma_t$. From now on, l_t in (18) will be denoted as l_t^{PAST}. We employ for PA densities the method of constrained maximum likelihood (CML) so as to guarantee the positivity of $\psi(\cdot)$ in (19). Indeed, we maximize $LL(\boldsymbol{\Upsilon})=\sum_{t=1}^T l_t^{PA}$ subject to $\boldsymbol{\Gamma}(v,\boldsymbol{\theta})<0$, which denotes a system of two nonlinear inequations as functions of $(v,\boldsymbol{\theta})$ according to (30) and (31) in BPZ (2015). Note that for each value of v, both $\boldsymbol{\Gamma}(v,\boldsymbol{\theta})<0$ and $\boldsymbol{\Lambda}(v)$ in section 2.4 become the same.

4 Empirical application

First, we start with the descriptive statistics analysis of the daily return series. Second, we estimate the return unconditional distribution for each series under different densities. Indeed, in this stage we are interested in the PAST relative performance for fitting the distribution tails. Third, we analyze the fit of the conditional distribution considering a model from the GARCH family under different densities for the standardized return series. Finally, we implement an exhaustive backtesting procedure to compare the distributions' OOS performance for predicting VaR and ES.

4.1 Data and descriptive statistics

We use the log returns computed as $r_t = 100 \ln(P_t/P_{t-1})$ from samples of daily closing prices $\{P_t\}_{t=1}^{T}$ of (1) exchange rates (FX): yen to euro (JAP-EU), yen to dollar (JAP-US), dollar to British pound (US-UK) and Swiss franc to dollar (SWI-US); (2) stock indexes: FTSE 100, CAC 40 and AEX; (3) Commodity indices: Goldman Sachs Standard & Poors (S&P) commodity index (GSCITOT), GSGCTOT S&P Gold index (GOLD) and GSBRSPT S&P Brent Crude Oil (BRENT). The data, downloaded from Datastream, cover the period from December 9, 2007 to April 4, 2021 for a total of $T = 3,500$ observations. Table 1 provides the return descriptive statistics. All series present negative skewness ranging from -0.23 (CAC) to -1.21 (SWI-US), and high kurtosis ranging from 8.17 (JAP-US) to 41.33 (SWI-US). The non-reported Jarque-Bera test null of normality is rejected in all cases motivating the use of alternative distributions to the Gaussian for modeling returns.

[Table 1]

4.2 Unconditional distribution estimation

We showcase the estimation of the models for the return series standardized by their sample means and standard deviations. Table 2 presents the parameter estimates of the PAST density as well as alternative pdfs considered for robustness comparison purposes. Besides, to assess the stability in the PAST estimation, we also consider PAST with fixed $\nu = 10, 15, 20$ (i.e., PAST10/15/20). The estimation is carried out using the ML method for both T and sk-T densities, and CML for all PA densities.

The parameter estimates θ_3, θ_4 (PA densities), λ (sk-T) and ν (T, sk-T and PAST) are all statistically significant for stock indices, GSCITOT, GOLD and BRENT series, indicating skewness and leptokurtosis. For FX returns the results are mixed: (1) both ν and θ_4 are statistically significant for all series indicating leptokurtosis; (2) the sk-T asymmetry parameter, λ, and θ_3 from the rest of models are not statistically significant for JAP-US indicating that the unconditional distribution of this series is symmetric; (3) for JAP-EU and US-UK the parameter estimates indicate milder or no skewness. In particular, for JAP-EU, the λ parameter is significant at 5 per cent; and θ_3 is significant at 5 per cent (GC), 10 per cent (PAST, PAST15 and PAST20) or not significant (PAL, PAHS, PAST10). For US-UK series, λ is significant at 5 per

cent level, but θ_3 is not at any reasonable significance level in any of the other densities. (4) Finally, for SWI-US, λ and θ_3 are both statistically significant at 5 per cent level, except in PAHS and PAST.

[Table 2]

For the goodness-of-fit (GoF) comparison of nested pdfs, we apply the Akaike Information Criterion (AIC). We find that sk-T performs better than T for all series except for JAP-US, for which the λ parameter is not significant indicating that the distribution is symmetric. This finding shows the importance of capturing skewness. The three-parameter PAST log-likelihood values (not reported) are always larger than those of PAST with restricted ν, but differences are very small in most cases to turn into better fit according to the AIC. Only for SWI-US and AEX, the three-parameter PAST provides superior fit. The PAST20 only does it better for JAP-EU while the PAST10 for US-UK. For the rest of the six series the PAST15 performs better.

Table 3 analyzes the behavior of both tails under the alternative distributions fitted to the standardized return series, with parameters in Table 2, by counting those observations that fall outside the specific interval $[-3, 3]$.[6] Note that under the Normal density, the probability of observations outside the previous interval is 0.27%, which is at least three times lower than the one observed in all series. Each cell in the table exhibits two values: the 'lower' or 'left' tail corresponding to the number of observations out of $T = 3,500$ lower than -3, and the 'upper' or 'right' tail (in parentheses) as the number of observations higher than 3. The first row with title 'empirical' shows the same information respecting the return series empirical distributions. We obtain the following results. First, all series exhibit longer left tails than right ones according to all fitted skewed densities such as sk-T and PA ones. This empirical evidence is in line with the negative skewness values displayed in Table 1 for all series. Second, the asymmetry effect on the distribution tail behavior can be seen if we compare the symmetric T (that is more suitable than N due to the high kurtosis levels shown in Table 1) with the other densities. All series show longer left tails under both sk-T and PA because of the negative skewness. The only exceptions correspond to JAP-EU and JAP-US under the PAST10 with a total of observations of 21 and 20, respectively. These numbers are lower than 22 under T in both series. Third, the same previous analysis applied to the right tail concludes that in most cases, as expected, there are now more observations under the T distribution. Note that most exceptions occur under the PAST20 that shows longer right tails than T. Fourth, if we compare sk-T and PAST we see that PAST exhibit longer tails than sk-T in most cases, except for SWI-US and, to a lesser extent, for US-UK. Fifth, the sk-T underestimates both tails of the empirical distribution in seven cases out of the ten. Note, for instance, the SWI-US case where sk-T (also PAL, PAHS and GC) overestimates both empirical tails, nevertheless PAST only does it for the left tail.

[Table 3]

Finally, a more in-depth analysis of the tail behavior, respecting the previous one in Table 3, is displayed in Figure 5. This figure presents a comparison of the theoretical quantiles of a density with the sample

[6] See Tolikas (2014) for a rather similar analysis.

standardized returns ones for both distribution tails. The tails are measured as the range of quantiles from 0.001 to 0.05 for the left tail, and 0.95 to 0.999 for the right tail. Specifically, a total of fifty equally-spaced quantiles in each tail. To shorten, we only select four densities as the most representative ones: sk-T, PAL, PAHS and PAST. As an example, this figure only displays four return series: JAP-US, GSCITOT, GOLD and BRENT. Each series analysis is exhibited in both panel A (lower tail) and panel B (upper tail), where each panel contains four graphs (each for a different density). First, the results do not throw a clear-cut better model between sk-T and PAST for both tails in most series except for JAP-US, where sk-T seems to fit the lower or left tail better than PAST while a similar fit occurs for the upper or right tail. Second, PAL, PAHS and PAST make a similar performance in most cases. For BRENT, the PAST performs slightly better than PAL and PAHS for the left tail.

[Figure 5]

4.3 Conditional distribution estimation

Next, we obtain the estimation of models for the return series conditional distribution. Table 4 presents the parameter estimates of the PAST density with the GJR conditional variance (GJR-PAST). First, the mean parameter, μ, is not statistically significant for any return series. Second, the GJR equation parameter estimates indicate presence of persistence in the conditional variance as well as asymmetric response of volatility to positive and negative shocks. These estimates are very similar to the quasi-ML estimates (QMLE), so they are not reported to save space. Third, the θ_3 parameter estimates show presence of negative skewness, which is statistically significant at least at 5 per cent level for all series except for US-UK. Fourth, the estimates for θ_4 and ν indicate excess kurtosis in the distribution of all return series. Fifth, the unconditional standard deviations (std) implied by the GJR-PAST model for returns, i.e. $\sigma = \sqrt{\alpha_0/(1 - E(c_t))}$, are very close to the sample ones. For instance, the estimation of σ is equal to 0.62 and the sample std is 0.63 for JAP-US.

[Table 4]

Table 5 provides the parameter estimates for the alternative densities. As for the PAST, the GJR parameter estimates in all cases are very similar to the non-reported QMLE ones. The parameter estimates show that after accounting for GARCH effects there is still statistically significant skewness and kurtosis in all cases. The discussion of the results is close to that of Table 2. In regard to the relative GoF, the sk-T performs better than the T for all series except for the JAP-US as for the unconditional distribution estimation. Among the different PAST models, we conclude the following. First, only for JAP-EU and US-UK, the three-parameter PAST provides superior fit. Second, PAST10 does it better for SWIS-US and BRENT series, while PAST20 for FTSE and AEX. Finally, PAST15 performs better for the remaining four series.

[Table 5]

Summing up, the analysis in this section shows that the PAST densities provide improvements in some cases with respect to the alternative densities considered. A cross comparison of all models, including non-nested ones, as regards forecasting VaR and ES is performed through a backtesting analysis next.

4.4 Backtesting analysis

We take the first T-N observations for the first in-sample window and an OOS period of length $N = 1,000$ using a daily constant-sized rolling window. We adopt a two-stage estimation method to each window as can be seen, among others, in Zhu and Galbraith (2011). The mean and GJR parameters are estimated by QML, then each density parameters are obtained by ML using the standardized residuals, z_t, from the first stage. We evaluate the forecasting performance for the returns' lower tail that is, for instance, of particular interest to pension fund managers concerned with the probabilities of losing a large part of investment portfolio value in a single day. We study the OOS performance, through VaR and ES backtesting, under alternative distributions of z_t.

4.4.1 Backtesting tests

Consider a nominal coverage rate α, the one-day conditional VaR is given by

$$VaR_t(\alpha) = \kappa_{0,t} + \kappa_{1,t} Q^{-1}(\alpha), \tag{20}$$

where $\kappa_{0,t} = \mu + a\sigma_t$ and $\kappa_{1,t} = b\sigma_t$. Let $h_t(\alpha) = \mathcal{I}(r_t < VaR_t(\alpha))$ denote the violation or hit variable. We are interested in checking whether the centered violations $\{h_t(\alpha) - \alpha\}_{t=1}^{\infty}$ follow a martingale difference sequence (MDS), which implies zero mean property and no correlation. Testing MDS leads to both the unconditional backtest (or unconditional coverage test) and conditional backtest (or independence test). The null hypothesis for the unconditional backtest, $H_{0,U} : E[h_t(\alpha)] = \alpha$, corresponds to the well-known test statistics by Kupiec (1995):

$$U_{VaR}(\alpha) = \frac{\sqrt{N}\left(\overline{h}(\alpha) - \alpha\right)}{\sqrt{\alpha(1-\alpha)}} \overset{a}{\sim} N(0,1), \tag{21}$$

where $\overline{h}(\alpha)$ is the sample average of $\left\{\widehat{h}_t(\alpha)\right\}_{t=1}^{N}$ such that $\widehat{h}_t(\alpha) = \mathcal{I}(\widehat{u}_t \leq \alpha)$ with $\widehat{u}_t$ as the estimation of $u_t = F(r_t | I_{t-1})$ where $F(\cdot | I_{t-1})$ denotes the conditional cdf for returns according to the pdf of z_t. For testing the conditional backtest null hypothesis, $H_{0,C} : E[h_t(\alpha) - \alpha | I_{t-1}] = 0$, we implement the approach by Escanciano and Olmo (2010) based on the Box-Pierce test statistic:

$$C(m) = N \sum_{i=1}^{m} \widehat{\rho}_j^2 \overset{a}{\sim} \chi_m^2, \tag{22}$$

which is asymptotically a chi-square distribution with m degrees of freedom such that $\hat{\rho}_j$ is the j-th lag of the sample autocorrelation defined as $\hat{\rho}_j = \hat{\gamma}_j / \hat{\gamma}_0$ where

$$\hat{\gamma}_j = \frac{1}{N-j} \sum_{t=1+j}^{N} \left(\widehat{h}_t(\alpha) - \alpha\right)\left(\widehat{h}_{t-j}(\alpha) - \alpha\right). \tag{23}$$

The unconditional and conditional ES backtests are the analogues to the above VaR ones. Du and Escanciano (2017) provide the ES backtest based on the notion of cumulative violations (CV), which accumulates the violations across the tail distribution and can be rewritten as

$$\mathcal{H}_t(\alpha) = \int_0^{\alpha} h_t(u)\, du = \frac{1}{\alpha}(\alpha - u_t)\mathcal{I}(u_t \leq \alpha). \tag{24}$$

Note that the equation (24) measures the distance of the returns from the corresponding α-quantile in (20) during the violations. It is shown that $\{\mathcal{H}_t(\alpha) - \alpha/2\}_{t=1}^{\infty}$ follows the MDS property. The null hypothesis for the unconditional backtest is $H_{0,U} : E[\mathcal{H}_t(\alpha)] = \alpha/2$ and the related test statistics is given by

$$U_{ES} = \frac{\sqrt{N}\left(\overline{\mathcal{H}}(\alpha) - \frac{\alpha}{2}\right)}{\sqrt{\alpha\left(\frac{1}{3} - \frac{\alpha}{4}\right)}} \overset{a}{\sim} N(0,1), \tag{25}$$

where $\overline{\mathcal{H}}(\alpha)$ is the mean of $\left\{\widehat{\mathcal{H}}_t(\alpha)\right\}_{t=1}^{N}$ such that $\widehat{\mathcal{H}}_t(\alpha) = \frac{1}{\alpha}(\alpha - \widehat{u}_t)\mathcal{I}(\widehat{u}_t \leq \alpha)$. The null for the conditional backtest hypothesis is $H_{0,C} : E[\mathcal{H}_t(\alpha)\,|I_{t-1}] = \alpha/2$ with the same test statistics in (22) such that

$$\hat{\gamma}_j = \frac{1}{N-j}\sum_{t=1+j}^{N}\left(\widehat{\mathcal{H}}_t(\alpha) - \frac{\alpha}{2}\right)\left(\widehat{\mathcal{H}}_{t-j}(\alpha) - \frac{\alpha}{2}\right). \tag{26}$$

4.4.2 Backtesting results

Following Deng and Qiu (2021), and references therein, a larger coverage level α for ES than VaR is selected to compare both risk measures. Specifically, we consider the following rule-of-thumb: the coverage level for ES is twice, or close to twice, than that of VaR. We focus on $\alpha = 2.5\%$ and $\alpha = 5\%$ for ES, corresponding roughly to $\alpha = 1\%$ and $\alpha = 2.5\%$ for VaR in a standard normal distribution.[7] Table 6 shows the results of the descriptive analysis of violations. First, all models perform better than the Normal for both VaR(1%) and ES(2.5%). Respecting VaR(2.5%) and ES(5%), there are very few exceptions for which the performance of some models is worse than that of the Normal, and primarily for the FX series: JAP-US and SWI-US. Second, most skewed density models do not perform worse than the T, with some exceptions mainly again found in the FX series. Indeed, there are no exceptions for any of the stock index series. Third, the GC (or PAST when $\nu \to \infty$) works better than sk-T for most series and coverage levels. Fourth, the PAST works similarly or better than the sk-T for stock and commodity indexes, whilst sk-T does it better than PAST for the FX series. Fifth, PAHS performs better than PAL for most series. Sixth, among the two-parameter PAST (i.e., PAST10/15/20) densities, PAST20 makes the best performance. Furthermore, PAST20 beats sk-T for all series (and coverage levels) except for US-UK. This result is in line with the already mentioned very good performance of GC since both densities tend to resemble each other for higher values of ν, see section 3.2

Table 6 also reports the significance at five percent level of both unconditional and conditional backtesting for VaR and ES (see superscripts u and c for the cases of rejecting the null hypotheses). For the unconditional

[7] For a discussion about the correspondence between coverage levels of ES and VaR, see Kerkhof and Melenberg (2004).

tests, the null is not accepted for most stock and commodity indexes under both N and T, whilst there is no rejection for FX series for these symmetric distributions. Respecting the conditional tests, there are many more rejections than in the case of unconditional ones independently of the densities. For the series JAP-EU, JAP-US, GOLD and BRENT there are hardly any rejections of the conditional null.

Finally, the previous VaR results are reinforced by the magnitude of exceptions for VaR measured through the quadratic loss (QL) function, see López (1999). The QL incorporates the exception magnitude and so, it provides useful information to discriminate among similar models according to the unconditional coverage test, i.e. $QL_t(\alpha) = (r_t - VaR_t(\alpha))^2 \times h_t(\alpha)$. We are interested in the sample average of QL (AQL) for the OOS period of N observations. Table 6 only exhibits both the first and second best models according to the AQL measure and denoted, respectively, with the symbols ♣ and ♦. [8] First, we find that the Normal renders the highest AQL values. Second, for SWI-US, CAC, AEX, GOLD and BRENT either PAST or PAST20 provides the lowest AQL at 1%, while PAHS performs better for US-UK, FTSE and GSCITOT at 1%. Third, GC provides the lowest AQL in most cases at 2.5%. Note that the second best model is provided by the PAST family, except for PAST10, in most series at both 1% and 2.5% levels.

[Table 6]

5 Conclusions

We present a polynomial expansion of the standardized Student-t distribution, referred to as PAST. The density belongs to the polynomially adjusted (PA) class of Bagnato et al. (2015), and it is a generalization of the Gram-Charlier (GC) density in Jondeau and Rockinger (2001). The two parameters in the polynomial expansion are by construction the skewness and excess kurtosis of this new density. We derive its parametric properties including the moments, the distribution function and the skewness and kurtosis frontiers (SKF) for which the density is well-defined. We show how the PAST's SKF enlarges that of the GC.

The performance of the PAST is tested through an empirical application to different types of asset returns: exchange rates, stock indexes and commodities. We consider several distributions for comparison purposes, including: Normal, Student-t, Hansen's skewed T, GC, PA logistic and PA hyperbolic secant. For robustness checks we also consider two-parameter PAST densities where the degrees of freedom are fixed to 10, 15 and 20. We find that the estimated PAST features flexibility to capture both skewness and high levels of kurtosis for both the unconditional and conditional distributions of the return series. Our in-depth in-sample analysis shows that the PAST density is capable to provide improvements respecting the alternative distributions. A more general analysis based on backtesting VaR and ES shows that the PAST performance can beat the alternative densities.

[8] To save space, the AQL values and all previous backtesting test statistics are not reported here, but they are available upon request.

Acknowledgements

Financial support from the Spanish Ministry of Economy and Competitiveness through grant ECO2017-87069-P is gratefully acknowledged by the first author.

Appendix

Proof of Corollary 1. The cdf of the standardized ST distribution in (14) is obtained by using previously the integral of a symmetric region, i.e.

$$\int_{-t}^{t} f(u;v)\,du = c\int_{-t}^{t}\left(1+\frac{u^2}{v-2}\right)^{-m} du = 2c\int_{0}^{t}\left(1+\frac{u^2}{v-2}\right)^{-m} du,$$

such that $t > 0$, $m = (v+1)/2$ and $c = (v-2)^{-1/2} B(v/2,1/2)^{-1}$. If we make the substitution $y^{-1} = 1+\frac{u^2}{v-2}$, we obtain

$$u = (v-2)^{1/2}(1-y)^{1/2} y^{-1/2}, \quad du = -\frac{1}{2}(v-2)^{1/2} y^{-3/2}(1-y)^{-1/2}\,dy. \tag{27}$$

Then,

$$\begin{aligned} 2c\int_{0}^{t}\left(1+\frac{u^2}{v-2}\right)^{-m} du &= B(v/2,1/2)^{-1}\int_{\theta(t)}^{1} y^{m-3/2}(1-y)^{-1/2}\,dy \\ &= B(v/2,1/2)^{-1}\int_{\theta(t)}^{1} y^{v/2-1}(1-y)^{-1/2}\,dy \\ &= B(v/2,1/2)^{-1}\left[B(v/2,1/2) - B(\theta(t);v/2,1/2)\right] \\ &= 1 - I_{\theta(t)}(v/2,1/2) = I_{1-\theta(t)}(1/2,v/2), \end{aligned}$$

where $\theta(t) = \frac{v-2}{t^2+v-2}$ and $1-\theta(t) = \frac{t^2}{t^2+v-2}$. Hence, $F(t) = \frac{1}{2}+\frac{1}{2}I_{\eta(t)}(1/2,v/2)$ with $\eta(t) = 1-\theta(t)$ and $F(-t) = 1-F(t) = \frac{1}{2}-\frac{1}{2}I_{\eta(t)}(1/2,v/2)$. Finally, we have $F(x) = \frac{1}{2}+\frac{1}{2}\,sgn(x)\,I_{\eta(x)}(1/2,v/2)$ where $x\in\mathbb{R}$. ■

It should be noted that the following proofs use the previous change of variable in (27).

Proof of Proposition 1. First, an alternative expression for the even-moments in (13), that is useful for the proof, is given by

$$\begin{aligned} m_{2k}(v) &= \int_{-\infty}^{+\infty} u^{2k} f(u;v)\,du = 2c\int_{0}^{+\infty} u^{2k}\left(1+\frac{u^2}{v-2}\right)^{-m} du \\ &= (v-2)^k B(v/2,1/2)^{-1}\int_{0}^{1}(1-y)^{k-1/2} y^{m-k-3/2} dy \\ &= (v-2)^k B(v/2,1/2)^{-1} B(v/2-k,k+1/2). \end{aligned} \tag{28}$$

Second, the even-moment $\xi_{2k}(t;v) = \int_{-\infty}^{t} u^{2k} f(u;v)\,du$ with $t > 0$ in (15) is obtained as

$$\begin{aligned}
\xi_{2k}(t;v) &= \frac{m_{2k}}{2} + c\int_0^t u^{2k}\left(1+\frac{u^2}{v-2}\right)^{-m} du \\
&= \frac{m_{2k}}{2} + \frac{(v-2)^k}{2B(v/2,1/2)}\int_{\theta(t)}^{1} y^{m-k-3/2}(1-y)^{k-1/2}\,dy \\
&= \frac{m_{2k}}{2} + \frac{(v-2)^k}{2B(v/2,1/2)}\left[B(v/2-k,k+1/2) - B(\theta(t);v/2-k,k+1/2)\right] \\
&= \frac{m_{2k}}{2} + \frac{(v-2)^k}{2}\frac{B(v/2-k,k+1/2)}{B(v/2,1/2)}\left[1 - I_{\theta(t)}(v/2-k,k+1/2)\right] \\
&= \frac{m_{2k}}{2} + \frac{(v-2)^k}{2}\frac{B(v/2-k,k+1/2)}{B(v/2,1/2)} I_{\eta(t)}(k+1/2,v/2-k). \qquad (29)
\end{aligned}$$

Note that $\lim_{t\to+\infty}\eta(t) = 1$ in (29) and hence, $\lim_{t\to+\infty} I_{\eta(t)}(k+1/2,v/2-k) = I_1(k+1/2,v/2-k) = 1$. If we consider (28), then it is shown that $\lim_{t\to+\infty}\xi_{2k}(t;v) = m_{2k}$. Finally, $\xi_{2k}(-t;v)$ with $t > 0$ can be expressed as

$$\xi_{2k}(-t;v) = m_{2k} - \xi_{2k}(t;v). \qquad (30)$$

It is verified that $\lim_{t\to+\infty}\xi_{2k}(-t;v) = m_{2k} - m_{2k} = 0$. If we consider both (29) and (30), we have $\xi_{2k}(x;v)$ where $x \in \mathbb{R}$ in (15). ■

Proof of Proposition 2. The odd-moment $\xi_{2k+1}(t;v) = \int_{-\infty}^{t} u^{2k+1} f(u;v)\,du$ with $t > 0$ in (16) can be rewritten as

$$\xi_{2k+1}(t;v) = -c\int_0^{+\infty} u^{2k+1}\left(1+\frac{u^2}{v-2}\right)^{-m} du + c\int_0^t u^{2k+1}\left(1+\frac{u^2}{v-2}\right)^{-m} du. \qquad (31)$$

To shorten, let $q = 2k+1$, $\varphi = \frac{(v-2)^{q/2}}{2B(v/2,1/2)}$, $a_1 = m - (q+1)/2$ and $a_2 = (q+1)/2$. Then,

$$c\int_0^{+\infty} u^q\left(1+\frac{u^2}{v-2}\right)^{-m} du = \varphi\int_0^1 (1-y)^{(q-1)/2} y^{m-(q+3)/2} dy = \varphi B(a_1,a_2), \qquad (32)$$

and

$$\begin{aligned}
c\int_0^t u^q\left(1+\frac{u^2}{v-2}\right)^{-m} du &= \varphi\int_{\theta(t)}^1 (1-y)^{(q-1)/2} y^{m-(q+3)/2} dy \\
&= \varphi\left[B(a_1,a_2) - B(\theta(t);a_1,a_2)\right] \\
&= \varphi B(a_1,a_2)\left[1 - I_{\theta(t)}(a_1,a_2)\right] \\
&= \varphi B(a_1,a_2) I_{\eta(t)}(a_2,a_1). \qquad (33)
\end{aligned}$$

By plugging (32) and (33) into (31), we have

$$\begin{aligned}
\xi_{2k+1}(t;v) &= \varphi B(a_1,a_2)\left[I_{\eta(t)}(a_2,a_1) - 1\right] \\
&= (v-2)^{k+\frac{1}{2}}\frac{B\left(\frac{v-1}{2}-k,k+1\right)}{2B\left(\frac{v}{2},\frac{1}{2}\right)}\left[I_{\eta(t)}\left(k+1,\frac{v-1}{2}-k\right) - 1\right]. \qquad (34)
\end{aligned}$$

Finally, $\xi_{2k+1}(-t;v)$ with $t>0$ can be expressed as

$$\begin{aligned}\xi_{2k+1}(-t;v) &= -\int_t^{+\infty} u^{2k+1} f(u;v)\,du = -\varphi\int_0^{\theta(t)} (1-y)^{(q-1)/2} y^{m-(q+3)/2} dy \\ &= -\varphi B(\theta(t);a_1,a_2) = \cdots = \varphi B(a_1,a_2)\left[1-I_{\eta(t)}(a_2,a_1)\right]. \end{aligned} \tag{35}$$

Note that $\lim_{t\to+\infty}\eta(t)=1$ in (35), hence $\lim_{t\to+\infty} I_{\eta(t)}\left(k+1,\frac{v-1}{2}-k\right)=1$ and $\lim_{t\to+\infty}\xi_{2k+1}(-t;v)=0$. Finally, if we consider both (34) and (35), we have $\xi_{2k+1}(x;v)$ where $x\in\mathbb{R}$ in (16). ■

References

[1] Bagnato, L., Potì, V., Zoia, M.G., 2015. The role of orthogonal polynomials in adjusting hyperbolic secant and logistic distributions to analyse financial asset returns. *Statistical Papers* 56(4), 1205-1234.

[2] Bali, T. G., Mo, H.,Tang, Y., 2008. The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking & Finance*, 32(2), 269-282.

[3] Chihara, T.S., 1978. An introduction to orthogonal polynomials, *Gordon & Breach, New York.*

[4] Corrado, C.J., Su, T., 1996. Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial research* 19(2), 175-192.

[5] Dendramis, Y., Spungin, G.E. and Tzavalis, E., 2014. Forecasting VaR models under different volatility processes and distributions of return innovations. *Journal of Forecasting*, 33(7), 515-531.

[6] Deng, K., Qiu, J., 2021. Backtesting expected shortfall and beyond. *Quantitative Finance*, 1-17.

[7] Del Brio, E.B., Perote, J., 2012. Gram–Charlier densities: Maximum likelihood versus the method of moments. *Insurance: Mathematics and Economics*, 51(3), 531-537.

[8] Du, Z., and Escanciano, J. C., 2017. Backtesting expected shortfall: accounting for tail risk. *Management Science* 63 (4), 940-958.

[9] Engle, R.F., Ng, V.K, 1993. Measuring and testing the impact of news on volatility. *Journal of Finance* 48(5), 1749-1778.

[10] Escanciano, J. C., Olmo, J., 2010. Backtesting parametric value-at-risk with estimation risk. *Journal of Business and Economic Statistics* 28 (1), 36-51.

[11] Feunou, B., Jahan-Parvar, M.R. and Tédongap, R., 2016. Which parametric model for conditional skewness?. *The European Journal of Finance* 22(13),.1237-1271.

[12] Gallant, A.R., Nychka, D.W., 1987. Seminonparametric maximum likelihood estimation. *Econometrica* 55, 363-390.

[13] Glosten, R.T., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48(5), 1779-1801.

[14] Hansen, B.E., 1994. Autoregressive conditional density estimation. *International Economic Review* 35(3), 705-730.

[15] Harvey, A. and Sucarrat, G., 2014. EGARCH models with fat tails, skewness and leverage. *Computational Statistics & Data Analysis* 76, 320-338.

[16] He, C., Teräsvirta, T., 1999. Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173-192.

[17] Kerkhof, J., and Melenberg, B., 2004. Backtesting for risk-based regulatory capital. *Journal of Banking & Finance* 28(8), 1845-1865.

[18] Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic dynamics and Control* 27(10), 1699-1737.

[19] Jondeau, E., Rockinger, M., 2001. Gram-Charlier densities. *Journal of Economic Dynamics and Control* 25(10), 1457-1483.

[20] Komunjer, I., 2007. Asymmetric power distribution: Theory and applications to risk measurement. *Journal of Applied Econometrics* 22(5), 891-921.

[21] Kumar P, H. and Patil S, B., 2016. Volatility Forecasting–A Performance Measure of Garch Techniques With Different Distribution Models. *International Journal of Soft Computing, Mathematics and Control* 5(2/3).

[22] Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 2, 174-184.

[23] León, A., Mencía, J., Sentana, E., 2009. Parametric properties of semi-nonparametric distribution, with applications to option valuation. *Journal of Business & Economic Statistics* 27(2), 176-192.

[24] León, Á. and Ñíguez, T.M., 2020. Modeling asset returns under time-varying semi-nonparametric distributions. *Journal of Banking & Finance* 118, 105870.

[25] León, A., Moreno, M., 2017. One-sided performance measures under Gram-Charlier distributions. *Journal of Banking & Finance* 74, 38-50.

[26] Liu, Y., Li, J.S.H. and Ng, A.C.Y., 2015. Option pricing under GARCH models with Hansen's skewed-t distributed innovations. *North American Journal of Economics and Finance* 31, 108-125.

[27] Lopez, J.A., 1999. Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review* 2, 3-17.

[28] Nicolussi, F., Zoia, M.G., 2020. Gram–Charlier-Like Expansions of the Convoluted Hyperbolic-Secant Density. *Journal of Statistical Theory and Practice* 14(1), 15.

[29] Peters, J.P., 2001. Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (Skewed) Student-t densities. University of Liege, 3, 19-34.

[30] Schlögl, E., 2013. Option pricing where the underlying assets follow a Gram/Charlier density of arbitrary order. *Journal of Economic Dynamics and Control* 37(3), 611-632.

[31] Thiele, S., 2020. Modeling the conditional distribution of financial returns with asymmetric tails. *Journal of Applied Econometrics*, 35(1), 46-60.

[32] Tolikas, K., 2014. Unexpected tails in risk measurement: Some international evidence. *Journal of Banking & Finance* 40, 476-493.

[33] Wilhelmsson, A., 2006. GARCH forecasting performance under different distribution assumptions. *Journal of Forecasting* 25(8), 561-578.

[34] Zhu, D. and Galbraith, J.W., 2011. Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions. *Journal of Empirical Finance* 18, 765-778.

[35] Zoia, M.G., Biffi, P. and Nicolussi, F., 2018. Value at risk and expected shortfall based on Gram-Charlier-like expansions. *Journal of Banking & Finance* 93, 92-104.

Tables

Table 1: Summary statistics for daily percent log returns

	JAP-EU	JAP-US	US-UK	SWI-US	FTSE	CAC	AEX	GSCITOT	GOLD	BRENT
Mean	-0.01	0.00	-0.01	-0.01	-0.01	0.00	0.00	-0.03	0.02	-0.01
Median	0.01	0.00	0.00	0.01	0.04	0.04	0.06	0.00	0.01	0.00
Std. dev.	0.75	0.63	0.64	0.66	1.49	1.67	1.55	1.49	1.14	2.29
Min	-6.79	-4.61	-8.31	-11.42	-14.21	-14.85	-13.13	-12.52	-9.81	-26.83
Max	4.84	3.71	4.47	8.47	12.22	12.14	12.32	7.62	8.59	19.08
Skewness	-0.40	-0.25	-0.61	-1.21	-0.41	-0.23	-0.31	-0.59	-0.24	-0.58
Kurtosis	9.66	8.17	14.73	41.33	15.02	11.64	13.05	9.28	9.18	16.31

This table presents the summary statistics for daily percent log return series of: yen to euro (JAP-EU); yen to dollar (JAP-US); dollar to British pound (US-UK); Swiss franc to dollar (SWI-US); FTSE 100, CAC 40 and AEX stock indexes; Goldman Sachs Standard & Poors (S&P) commodity index (GSCITOT); GSGCTOT S&P Gold index (GOLD); and GSBRSPT S&P Brent Crude Oil (BRENT). Sample: November 9, 2007 to April 4, 2021 ($T = 3,500$ observations).

Table 2: Unconditional distribution estimation for standardized returns

	JAP-EU	JAP-US	US-UK	SWI-US	FTSE	CAC	AEX	GSCITOT	GOLD	BRENT
T										
v	3.600***	3.689***	4.302***	3.483***	2.968***	3.113**	3.055***	3.385***	3.351***	3.108***
	(0.130)	(0.143)	(0.217)	(0.121)	(0.063)	(0.078)	(0.071)	(0.107)	(0.109)	(0.080)
AIC	2.6576	2.6741	2.6840	2.5639	2.5031	2.5644	2.5342	2.6350	2.6363	2.5586
sk-T										
λ	-0.054**	-0.022	-0.042**	-0.044**	-0.069***	-0.052***	-0.071***	-0.067***	-0.030**	-0.050***
	(0.016)	(0.016)	(0.019)	(0.017)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
v	3.591***	3.689***	4.293***	3.478***	2.964***	3.138***	3.057***	3.378***	3.358***	3.107***
	(0.128)	(0.143)	(0.216)	(0.120)	(0.062)	(0.078)	(0.071)	(0.107)	(0.110)	(0.080)
AIC	2.6555	2.6742	2.6833	2.5629	2.4984	2.5621	2.5293	2.6309	2.6359	2.5564
PAL										
θ_3	-0.157	-0.049	-0.152	-0.246**	-0.273**	-0.411***	-0.490***	-0.348***	-0.285***	-0.294***
	(0.107)	(0.107)	(0.112)	(0.123)	(0.110)	(0.103)	(0.108)	(0.102)	(0.105)	(0.107)
θ_4	3.825***	3.785***	3.367***	4.195***	6.002***	5.649***	5.782***	4.577***	4.342***	5.416***
	(0.420)	(0.419)	(0.435)	(0.434)	(0.372)	(0.379)	(0.371)	(0.402)	(0.405)	(0.390)
AIC	2.6739	2.6861	2.6869	2.5904	2.5434	2.5882	2.5617	2.6518	2.6573	2.5912
PAHS										
θ_3	-0.165	-0.048	-0.165	-0.249	-0.337**	-0.446***	-0.554***	-0.375***	-0.301**	-0.286**
	(0.138)	(0.133)	(0.138)	(0.163)	(0.139)	(0.129)	(0.133)	(0.124)	(0.130)	(0.133)
θ_4	3.541***	3.575***	3.350***	4.553***	6.953***	6.272***	6.816***	4.626***	4.351***	5.952***
	(0.619)	(0.631)	(0.648)	(0.657)	(0.568)	(0.580)	(0.564)	(0.622)	(0.615)	(0.607)
AIC	2.6613	2.6751	2.6849	2.5757	2.5234	2.5745	2.5445	2.6391	2.6417	2.5745
GC										
θ_3	-0.125**	-0.067	-0.067	-0.126**	-0.173***	-0.219***	-0.236***	-0.187***	-0.193***	-0.178***
	(0.055)	(0.057)	(0.057)	(0.059)	(0.058)	(0.056)	(0.057)	(0.057)	(0.057)	(0.058)
θ_4	1.979***	1.807***	1.467***	1.825***	2.474***	2.405***	2.379***	2.039***	1.981***	2.221***
	(0.102)	(0.101)	(0.107)	(0.107)	(0.094)	(0.096)	(0.097)	(0.099)	(0.098)	(0.097)
AIC	2.7144	2.7266	2.7606	2.7399	2.6390	2.6504	2.6495	2.6951	2.7019	2.6735

(Table 2 continued)

PAST										
θ_3	-0.142*	-0.056	-0.155	-0.238	-0.225***	-0.319***	-0.405***	-0.271***	-0.251**	-0.261***
	(0.074)	(0.082)	(0.116)	(0.152)	(0.087)	(0.079)	(0.096)	(0.080)	(0.078)	(0.086)
θ_4	3.436***	3.733***	8.087*	37.185***	6.205***	4.807***	6.560***	3.888***	3.823***	5.231***
	(0.609)	(0.701)	(4.458)	(11.154)	(1.288)	(0.706)	(1.494)	(0.636)	(0.657)	(1.121)
ν	18.447***	15.649***	10.184***	8.528***	13.820***	16.579***	12.941***	17.198***	16.972***	14.583***
	(4.207)	(2.781)	(1.369)	(0.165)	(1.919)	(2.365)	(1.636)	(2.911)	(2.983)	(2.285)
AIC	2.6785	2.6928	2.6871	2.5843	2.5524	2.5930	2.5715	2.6576	2.6645	2.5975
PAST10										
θ_3	-0.150	-0.044	-0.158	-0.252**	-0.249**	-0.390***	-0.468***	-0.334***	-0.276**	-0.285***
	(0.112)	(0.109)	(0.116)	(0.124)	(0.106)	(0.098)	(0.103)	(0.100)	(0.103)	(0.103)
θ_4	9.279***	9.294***	8.743***	10.494***	13.816***	13.019***	13.426***	10.872***	10.380***	12.677***
	(0.903)	(0.908)	(0.954)	(0.922)	(0.780)	(0.799)	(0.780)	(0.864)	(0.871)	(0.828)
AIC	2.6838	2.6959	2.6865	2.5860	2.5559	2.6020	2.5737	2.6648	2.6703	2.6023
PAST15										
θ_3	-0.146*	-0.055	-0.111	-0.222**	-0.221***	-0.332***	-0.371***	-0.286***	-0.258**	-0.259***
	(0.081)	(0.084)	(0.087)	(0.092)	(0.084)	(0.080)	(0.085)	(0.082)	(0.082)	(0.085)
θ_4	4.129***	3.915***	3.366***	4.049***	5.560***	5.358***	5.271***	4.487***	4.349***	5.044***
	(0.296)	(0.294)	(0.309)	(0.303)	(0.263)	(0.269)	(0.266)	(0.281)	(0.284)	(0.272)
AIC	2.6789	2.6923	2.6895	2.5936	2.5520	2.5927	2.5716	2.6574	2.6642	2.5969
PAST20										
θ_3	-0.141*	-0.060	-0.094	-0.196**	-0.209***	-0.299***	-0.324***	-0.256***	-0.242**	-0.238***
	(0.072)	(0.075)	(0.077)	(0.081)	(0.075)	(0.073)	(0.076)	(0.074)	(0.074)	(0.076)
θ_4	3.252***	3.025***	2.540***	3.078***	4.252***	4.114***	4.032***	3.446***	3.353***	3.832***
	(0.209)	(0.207)	(0.218)	(0.215)	(0.188)	(0.192)	(0.191)	(0.199)	(0.201)	(0.193)
AIC	2.6780	2.6931	2.6942	2.6009	2.5557	2.5933	2.5758	2.6574	2.6644	2.5993

This table presents ML estimates of density model parameters for standardized returns, z_t. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. *AIC* denotes Akaike Information Criterion. (***) indicates significance at 1% level; (**) at 5% level and (*) at 10% level. Sample: November 9, 2007 to April 4, 2021 ($T = 3,500$ observations).

Table 3: Number of standardized returns outside the interval $[-3, 3]$

	JAP-EU		JAP-US		US-UK		SWI-US		FTSE		CAC		AEX		GSCITOT		GOLD		BRENT	
empirical	23	(25)	27	(23)	21	(19)	15	(16)	31	(23)	32	(23)	34	(20)	32	(23)	29	(24)	31	(24)
N	4	(4)	4	(4)	4	(4)	4	(4)	4	(4)	4	(4)	4	(4)	4	(4)	4	(4)	4	(4)
T	22	(22)	22	(22)	20	(20)	23	(23)	23	(23)	23	(23)	23	(23)	23	(23)	23	(23)	23	(23)
sk-T	25	(18)	24	(20)	23	(17)	24	(19)	26	(18)	25	(19)	26	(18)	27	(18)	25	(20)	25	(19)
PAL	27	(20)	25	(23)	26	(20)	30	(19)	36	(23)	38	(20)	41	(19)	32	(18)	31	(19)	35	(21)
PAHS	24	(19)	22	(21)	24	(19)	28	(18)	32	(19)	34	(16)	35	(15)	31	(16)	28	(17)	30	(20)
GC	29	(23)	25	(23)	22	(19)	28	(22)	35	(29)	36	(27)	36	(26)	31	(23)	31	(22)	33	(26)
PAST	32	(26)	28	(24)	22	(17)	18	(11)	38	(27)	42	(26)	39	(21)	36	(23)	35	(23)	37	(26)
PAST10	21	(17)	20	(18)	21	(17)	25	(16)	27	(19)	30	(14)	31	(14)	27	(15)	26	(16)	27	(17)
PAST15	31	(24)	28	(25)	26	(21)	33	(20)	40	(29)	41	(26)	42	(25)	36	(22)	34	(22)	38	(26)
PAST20	33	(26)	29	(27)	27	(23)	33	(24)	42	(31)	43	(28)	43	(26)	37	(25)	36	(25)	39	(27)

This table exhibits the number of standardized observations z_t under each assumed distribution (in row), with parameters exhibited in Table 2, for the alternative return series (in column) that lie outside the interval $[-3, 3]$. The *left tail* is the number of observations lower than -3 and the *right tail* is the number of observations (in parentheses) higher than 3. The first row with name 'empirical' is related to the empirical distribution of the return series. All series have the same number of observations, $T = 3,500$.

Table 4: GJR-PAST model estimation results

	JAP-EU	JAP-US	US-UK	SWI-US	FTSE	CAC	AEX	GSCITOT	GOLD	BRENT
μ	0.003	0.003	0.002	-0.005	-0.010	-0.001	0.008	-0.014	0.023	-0.002
	(0.008)	(0.007)	(0.007)	(0.007)	(0.016)	(0.018)	(0.016)	(0.018)	(0.015)	(0.026)
α_0	0.002^{**}	0.002^{***}	0.003^{***}	0.001^{***}	0.020^{***}	0.018^{***}	0.017^{***}	0.011^{**}	0.004^{*}	0.033^{***}
	(0.001)	(0.001)	(0.001)	(0.0005)	(0.007)	(0.006)	(0.006)	(0.004)	(0.002)	(0.012)
β	0.957^{***}	0.953^{***}	0.951^{***}	0.964^{***}	0.901^{***}	0.912^{***}	0.911^{***}	0.941^{***}	0.962^{***}	0.924^{***}
	(0.009)	(0.007)	(0.008)	(0.005)	(0.022)	(0.015)	(0.018)	(0.008)	(0.006)	(0.011)
α_1^+	0.027^{***}	0.036^{***}	0.027^{***}	0.028^{***}	0.012	0.007	0.010	0.027^{***}	0.040^{***}	0.034^{***}
	(0.007)	(0.006)	(0.006)	(0.005)	(0.011)	(0.007)	(0.008)	(0.008)	(0.009)	(0.009)
α_1^-	0.051^{***}	0.045^{***}	0.053^{***}	0.036^{***}	0.164^{***}	0.159^{***}	0.151^{***}	0.080^{***}	0.027^{***}	0.107^{***}
	(0.011)	(0.008)	(0.009)	(0.007)	(0.036)	(0.028)	(0.032)	(0.012)	(0.006)	(0.017)
θ_3	-0.226^{***}	-0.192^{**}	-0.013	-0.304^{**}	-0.317^{***}	-0.371^{***}	-0.380^{***}	-0.441^{***}	-0.208^{***}	-0.548^{***}
	(0.055)	(0.088)	(0.089)	(0.121)	(0.068)	(0.083)	(0.072)	(0.106)	(0.090)	(0.120)
θ_4	1.290^{***}	2.388^{**}	0.728	28.814^{***}	1.406^{***}	1.763^{***}	1.163^{***}	2.012^{*}	2.699^{***}	12.227^{**}
	(0.259)	(1.026)	(1.040)	(10.273)	(0.273)	(0.578)	(0.219)	(1.060)	(0.670)	(4.768)
v	18.364^{***}	13.146^{***}	9.421^{***}	8.122^{***}	23.625^{***}	14.952^{***}	25.394^{**}	12.273^{***}	16.033^{***}	8.538^{***}
	(4.081)	(4.485)	(1.200)	(0.038)	(8.583)	(4.402)	(12.236)	(4.052)	(4.275)	(0.291)
σ	0.684	0.620	0.566	0.628	1.392	2.00	1.427	1.376	1.042	2.443
AIC	1.9127	1.6846	1.5616	1.7136	3.0525	3.3470	3.1547	3.2565	2.7952	4.0268

Model: $r_t = \mu + \varepsilon_t$, $\varepsilon_t = \sigma_t z_t$, $\quad \sigma_t^2 = \alpha_0 + \beta\sigma_{t-1}^2 + \alpha_1^+(\varepsilon_{t-1}^+)^2 + \alpha_1^-(\varepsilon_{t-1}^-)^2$, $z_t \sim i.i.d.$ PAST (0,1) distributed with parameters (v, θ_3, θ_4)

This table presents CML estimates of the GJR-PAST parameters for the return series in Table 1. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. σ denotes the unconditional GJR-PAST standard deviation, and AIC is Akaike Information Criterion. Finally, (***) indicates significance at 1% level; (**) at 5% level and (*) at 10% level. Sample: November 9, 2007 to April 4, 2021 ($T = 3,500$ observations).

Table 5: Alternative density (D) parameter estimates from GJR-D model

	JAP-EU	JAP-US	US-UK	SWI-US	FTSE	CAC	AEX	GSCITOT	GOLD	BRENT
T										
v	6.156***	4.926***	7.978***	5.909***	6.048***	5.971**	6.620***	5.754***	3.951***	5.235***
	(0.653)	(0.417)	(1.209)	(0.733)	(0.597)	(0.602)	(0.709)	(0.583)	(0.280)	(0.493)
AIC	1.9274	1.6299	1.6926	1.6172	3.0559	3.3507	3.1622	3.2589	2.7845	4.0265
sk-T										
λ	-0.063***	-0.025	-0.048**	-0.064***	-0.117***	-0.116***	-0.134***	-0.094***	-0.034*	-0.080***
	(0.021)	(0.022)	(0.022)	(0.022)	(0.021)	(0.023)	(0.022)	(0.020)	(0.020)	(0.020)
v	6.122***	4.937***	7.950***	5.902***	6.107***	6.175***	6.939***	5.984***	3.986***	5.357***
	(0.652)	(0.419)	(1.186)	(0.721)	(0.609)	(0.636)	(0.775)	(0.620)	(0.284)	(0.509)
AIC	1.9257	1.6301	1.6920	1.6155	3.0488	3.3440	3.1531	3.2544	2.7843	4.0232
PAL										
θ_3	-0.145	-0.116	-0.067	-0.150	-0.349***	-0.418***	-0.487***	-0.448***	-0.267**	-0.410***
	(0.104)	(0.113)	(0.092)	(0.110)	(0.088)	(0.094)	(0.101)	(0.095)	(0.106)	(0.099)
θ_4	0.980**	1.895***	0.284	1.048*	0.958**	1.237***	1.177***	1.159***	2.375***	1.204***
	(0.418)	(0.493)	(0.277)	(0.542)	(0.415)	(0.393)	(0.312)	(0.343)	(0.453)	(0.399)
AIC	1.9270	1.6306	1.6943	1.6266	3.0526	3.3461	3.1557	3.2536	2.7891	4.0232
PAHS										
θ_3	-0.053	-0.072	-0.083	-0.177	-0.329**	-0.425***	-0.358***	-0.415***	-0.280**	-0.447***
	(0.110)	(0.140)	(0.177)	(0.140)	(0.134)	(0.138)	(0.111)	(0.136)	(0.125)	(0.141)
θ_4	0.269	1.682**	0.145	0.906	0.610	1.248**	0.689**	1.123**	1.874***	1.287**
	(0.588)	(0.770)	(0.325)	(0.770)	(0.893)	(0.551)	(0.316)	(0.529)	(0.687)	(0.503)
AIC	1.9282	1.6296	1.7008	1.6279	3.0552	3.3504	3.1621	3.2554	2.7817	4.0221
GC										
θ_3	-0.172***	-0.074	-0.090	-0.148**	-0.266***	-0.282**	–0.319***	-0.267***	-0.150**	-0.233***
	(0.058)	(0.062)	(0.055)	(0.070)	(0.055)	(0.056)	(0.053)	(0.058)	(0.064)	(0.059)
θ_4	0.991***	1.421***	0.801***	1.489***	1.088***	1.053***	0.977***	1.014***	1.646***	1.134***
	(0.122)	(0.155)	(0.138)	(0.527)	(0.122)	(0.128)	(0.121)	(0.122)	(0.140)	(0.128)
AIC	1.9405	1.6551	1.7110	1.7298	3.0593	3.3611	3.1584	3.2706	2.8181	4.0477

(Table 5 continued)

PAST10										
θ_3	-0.190*	-0.134	-0.086	-0.167	-0.397***	-0.463***	-0.565***	-0.504***	-0.276**	-0.444***
	(0.113)	(0.119)	(0.107)	(0.114)	(0.092)	(0.095)	(0.100)	(0.096)	(0.108)	(0.102)
θ_4	3.248***	5.340***	1.117	2.948***	3.198***	3.731***	3.443***	3.461***	6.288***	3.637***
	(1.071)	(1.128)	(0.854)	(1.104)	(1.023)	(1.043)	(0.795)	(0.944)	(0.974)	(1.051)
AIC	1.9291	1.6340	1.6923	1.6176	3.0550	3.3471	3.1567	3.2562	2.7966	4.0265
PAST15										
θ_3	-0.213**	-0.107	-0.099	-0.147*	-0.351***	-0.370***	-0.447***	-0.398***	-0.216**	-0.329***
	(0.085)	(0.086)	(0.078)	(0.080)	(0.074)	(0.074)	(0.070)	(0.076)	(0.086)	(0.079)
θ_4	1.548***	2.410***	0.842***	1.591***	1.808***	1.758***	1.439***	1.582***	2.870***	1.853***
	(0.316)	(0.356)	(0.309)	(0.348)	(0.320)	(0.333)	(0.311)	(0.317)	(0.323)	(0.331)
AIC	1.9288	1.6338	1.6936	1.6211	3.0529	3.3464	3.1548	3.2561	2.7946	4.0274
PAST20										
θ_3	-0.205***	-0.097	-0.103	-0.139**	-0.327***	-0.340***	-0.403***	-0.355	-0.188**	-0.295***
	(0.075)	(0.076)	(0.069)	(0.071)	(0.067)	(0.067)	(0.064)	(0.070)	(0.077)	(0.072)
θ_4	1.289***	1.965***	0.818***	1.383***	1.507***	1.444***	1.244***	1.322	2.325***	1.531***
	(0.223)	(0.255)	(0.224)	(0.246)	(0.225)	(0.235)	(0.229)	(0.225)	(0.235)	(0.230)
AIC	1.9293	1.6347	1.6947	1.6245	3.0521	3.3468	3.1542	3.2569	2.7950	4.0286

This table presents either CML or ML estimates of different density (D) parameters from GJR-D model. GJR estimates are not presented to save space. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. *AIC* denotes Akaike Information Criterion. (***) indicates significance at 1% level; (**) at 5% level and (*) at 10% level. Sample: November 9, 2007 to April 4, 2021 ($T = 3,500$ observations).

Table 6: Descriptive analysis of violations and backtesting procedures

	VaR(1%)	ES(2.5%)	VaR(2.5%)	ES(5%)	VaR(1%)	ES(2.5%)	VaR(2.5%)	ES(5%)
			JAP-EU				JAP-US	
N	13	14.12	22	22.16	16	15.93	23	24.12
T	9	11.02	20	21.66	9	12.33	23	23.63
sk-T	9	9.82	19	19.71	8	12.11	23	23.31
PAL	9	10.49	20	20.94	10	12.66	23	23.92
PAHS	9	10.04	19	19.97	9	12.14	22	22.91
GC	$8^{c}_{\clubsuit}$	8.84	$18_{\clubsuit}$	17.29	$8_{\clubsuit}$	10.61	$18_{\clubsuit}$	20.40
PAST	9	10.56	21	21.32	12	13.26	26	24.77
PAST10	10	11.42	21	22.39	12	14.01	26	25.60
PAST15	9	10.05	20	20.56	9	12.37	24	24.17
PAST20	$8^{c}_{\spadesuit}$	9.53	$19_{\spadesuit}$	19.62	$8_{\spadesuit}$	11.64	$22_{\spadesuit}$	23.25
			US-UK				SWI-US	
Normal	13	15.44	29	28.04^{c}	12^{c}	14.43^{c}	24^{c}	22.94^{c}
T	8^{c}	11.91	29	26.45^{c}	8^{c}	10.59^{c}	23^{c}	22.39^{c}
sk-T	8^{c}	10.82	24	25.13^{c}	6^{c}	8.90^{c}	22^{c}	20.10^{c}
PAL	8^{c}	10.73	25	25.21^{c}	5^{c}	8.71^{c}	22^{c}	20.81^{c}
PAHS	$6_{\clubsuit}$	9.19	$21^{c}_{\clubsuit}$	23.80^{c}	5^{c}	8.45^{c}	22^{c}	20.36^{c}
GC	$8^{c}_{\spadesuit}$	10.30	$22^{c}_{\spadesuit}$	24.04^{c}	6^{c}	7.00^{c}	$15^{u,c}_{\clubsuit}$	$16.52^{u,c}$
PAST	8^{c}	11.70	28	26.42^{c}	9^{c}	12.68^{c}	25^{c}	25.23^{c}
PAST10	8^{c}	12.18	28	26.69^{c}	7^{c}	10.60^{c}	25^{c}	23.14^{c}
PAST15	8^{c}	12.13	28	26.62^{c}	$5^{c}_{\spadesuit}$	8.40^{c}	22^{c}	20.51^{c}
PAST20	8^{c}	11.71	28	26.20^{c}	$5^{c}_{\clubsuit}$	7.58^{c}	$20^{c}_{\spadesuit}$	19.27^{c}
			FTSE				CAC	
N	27^{u}	$25.84^{u,c}$	$38^{u,c}$	$34.93^{u,c}$	$24^{u,c}$	$22.79^{u,c}$	32^{c}	32.43^{c}
T	$17^{u,c}$	$21.09^{u,c}$	$37^{u,c}$	$33.10^{u,c}$	16	18.71^{u}	31^{c}	31.21^{c}
sk-T	16^{c}	17.79^{c}	29^{c}	29.31^{c}	15	16.26	27^{c}	27.57^{c}
PAL	16^{c}	18.01^{c}	32^{c}	30.30^{c}	14	15.84	27^{c}	28.21^{c}
PAHS	$13_{\clubsuit}$	16.49^{c}	$29^{c}_{\spadesuit}$	29.30^{c}	$14_{\spadesuit}$	15.55	27^{c}	27.93^{c}
GC	15^{c}	16.89^{c}	28	27.00^{c}	14	15.75	$24^{c}_{\clubsuit}$	25.05^{c}
PAST	$15^{c}_{\spadesuit}$	17.05^{c}	$29^{c}_{\clubsuit}$	28.80^{c}	14	16.30	27^{c}	28.69^{c}
PAST10	16^{c}	$19.33^{u,c}$	34^{c}	31.70^{c}	15	17.12	30^{c}	29.98^{c}
PAST15	15^{c}	17.59^{c}	32^{c}	29.92^{c}	14	15.96	27^{c}	28.13^{c}
PAST20	15^{c}	17.09^{c}	29^{c}	29.04^{c}	$14_{\clubsuit}$	15.59	$27^{c}_{\spadesuit}$	27.25^{c}

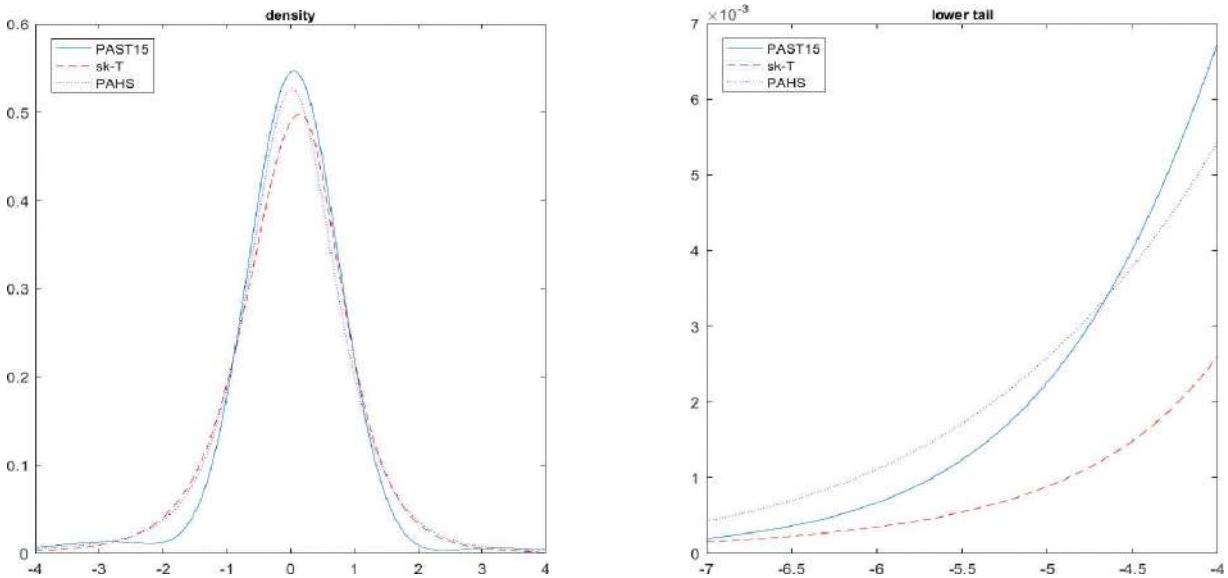
(Table 6 continued)

	VaR(1%)	ES(2.5%)	VaR(2.5%)	ES(5%)	VaR(1%)	ES(2.5%)	VaR(2.5%)	ES(5%)
	AEX				GSCITOT			
N	$22^{u,c}$	$25.10^{u,c}$	$38^{u,c}$	$35.29^{u,c}$	22^{u}	24.99^{u}	34	35.61^{u}
T	18^{u}	20.23^{u}	$37^{u,c}$	33.45^{u}	19^{u}	20.78^{u}	34	34.32^{u}
sk-T	14	16.37	30	29.08	18^{u}	18.20^{u}	31	30.40
PAL	12	15.73	31	29.61^{c}	$15_{\blacklozenge}$	16.74	30	29.89
PAHS	12	15.65	32	29.68	$15_{\clubsuit}$	16.11	30	29.93
GC	12	14.84	$26^{c}_{\clubsuit}$	26.16	17^{u}	17.51	$29_{\clubsuit}$	27.75
PAST	$11_{\clubsuit}$	15.02	$30_{\blacklozenge}$	27.72	17^{u}	18.00	32	31.01
PAST10	15	17.37	35^{c}	31.28^{c}	18^{u}	18.21^{u}	32	31.72
PAST15	11	15.54	30	29.04^{c}	17^{u}	17.41	30	30.07
PAST20	$11_{\blacklozenge}$	15.02	30	28.08	16	17.21	$29_{\blacklozenge}$	29.33
	GOLD				BRENT			
N	17^{u}	19.01^{u}	32	27.65	21^{u}	23.28^{u}	34	34.86^{u}
T	9	14.22	32	27.35	18^{u}	19.03^{u}	32	33.72^{u}
sk-T	7	13.33	30	26.39	17^{u}	16.90	29	30.36
PAL	7	12.20	28	25.41	14	16.09	28	29.94
PAHS	7	11.86	26	24.59	$14_{\blacklozenge}$	15.29	28	29.35
GC	7	9.56	$21_{\clubsuit}$	20.46	14	16.04	$21_{\clubsuit}$	26.59
PAST	$7_{\blacklozenge}$	11.22	26	24.52	16	16.89	30	30.70
PAST10	9	14.64	32	27.60	17^{u}	17.51	32	31.96
PAST15	7	11.68	27	25.14	14	16.24	29	30.11
PAST20	$7_{\clubsuit}$	10.65	$25_{\blacklozenge}$	23.87	$14_{\clubsuit}$	15.85	$28_{\blacklozenge}$	29.10

This table shows the violations for VaR and the cumulative violations in (24) from N, T, sk-T, **PAL**, **PAHS**, **PAST** and **PAST10/15/20** models. We also report the significance for (i) the VaR backtesting tests in (21) and (22) with $\hat{\gamma}_j$ in (23), and (ii) the ES backtesting tests in (25) and (22) with $\hat{\gamma}_j$ in (26). We set $m = 5$ in the Box-Pierce test statistic (22) for the two conditional backtests. The superscripts u and c indicate significance at five percent level for the unconditional and conditional backtests, respectively. The subscripts ♣ and ♦ indicate the first and second best models, respectively, for VaR by considering the magnitude of exceptions through the average of the quadratic losses, **AQL** (the lower **AQL**, the better). The OOS period covers from June 9, 2017 to April 8, 2021. Predictions: 1000.

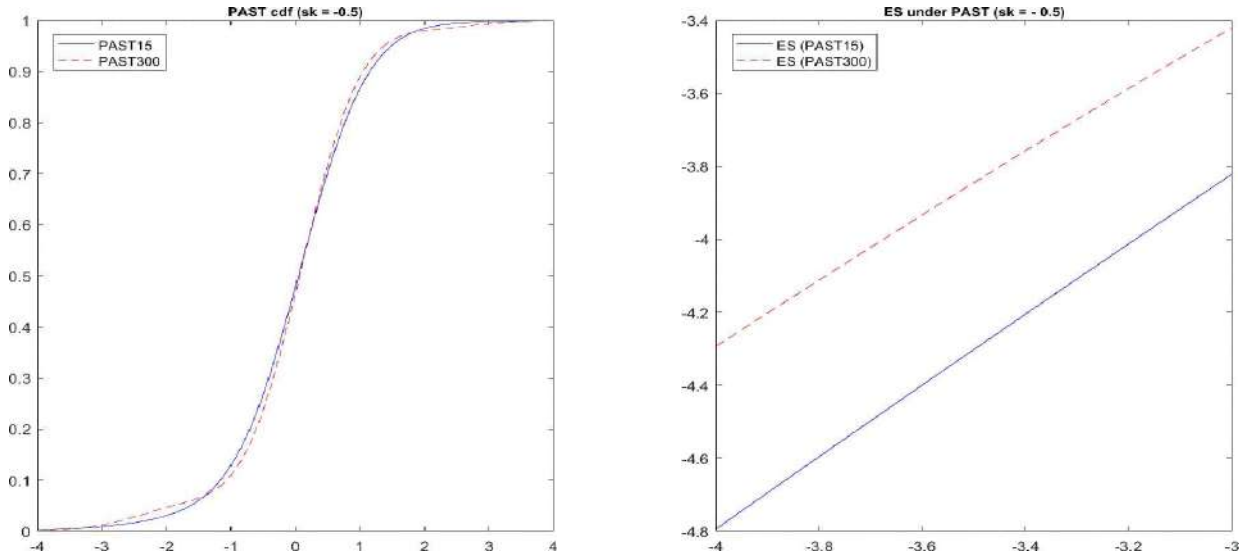
Figures

Figure 1: Distribution/tail shapes comparison



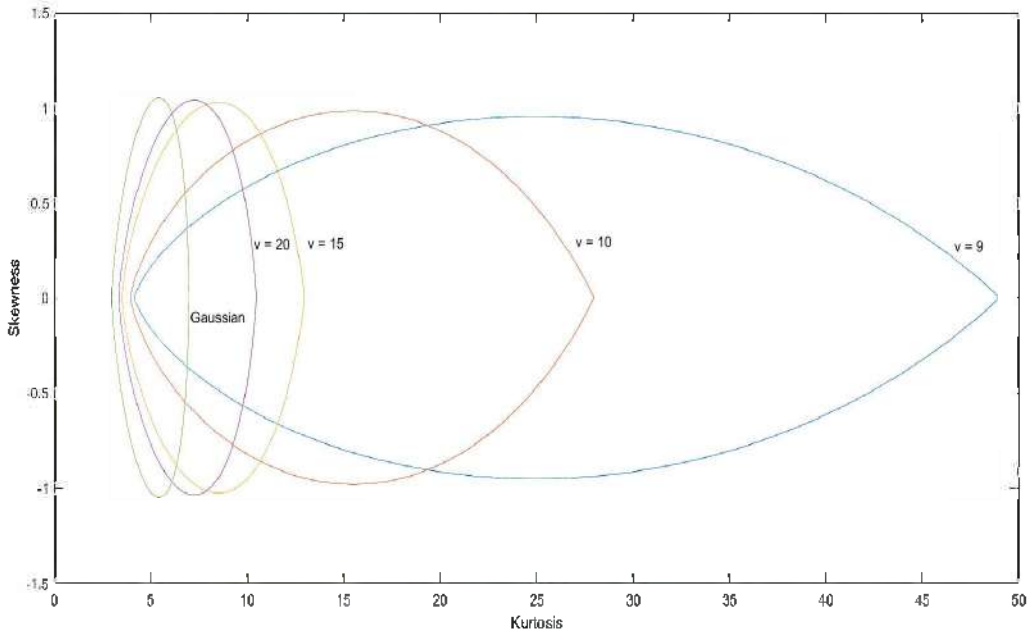
This figure provides a comparison of density and left-tail shapes of PAST15, sk-T and PAHS.

Figure 2: Cumulative distribution and expected shortfall



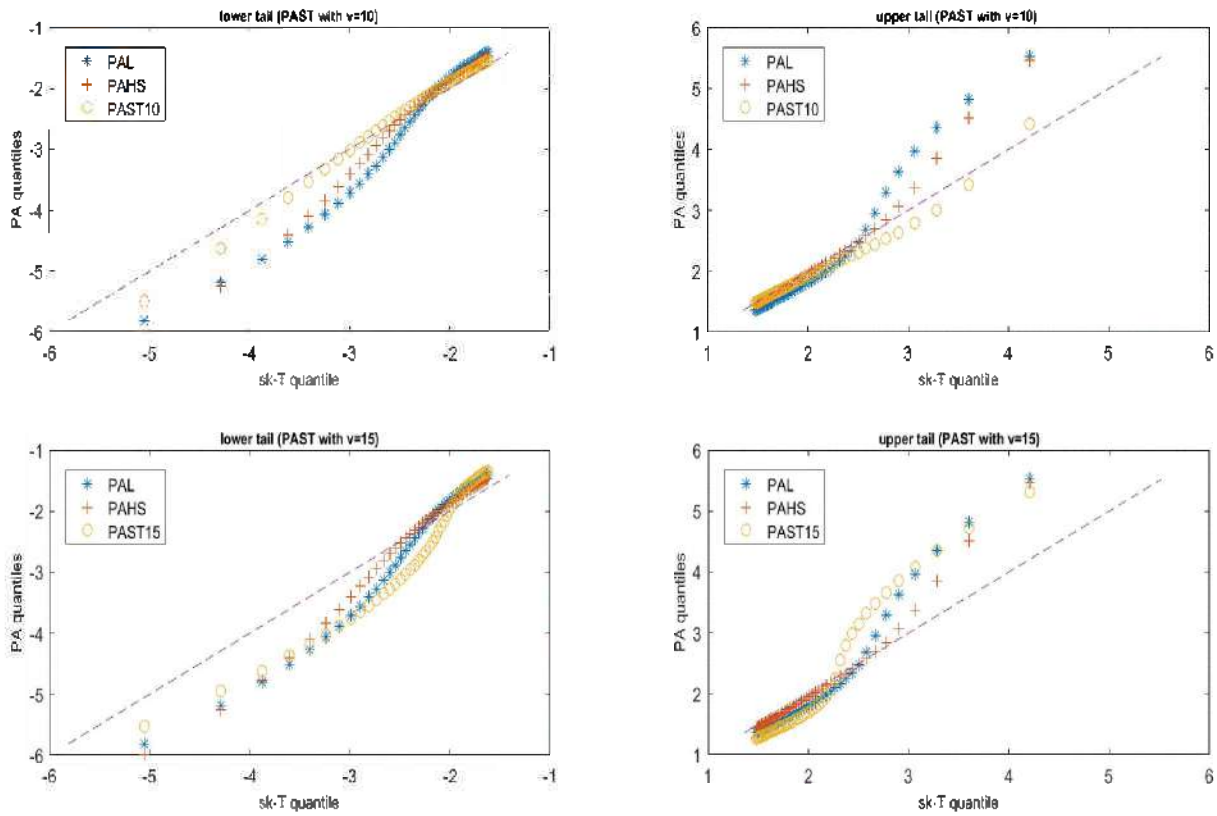
This figure plots cdf and ES of PAST15 and PAST300 (aprox. GC) pdfs.

Figure 3: PAST skewness-kurtosis frontiers



Skewness-kurtosis frontiers for PAST with $v = 9, 10, 15, 20$, and GC or PA Gaussian ($v \to \infty$).

Figure 4: Distribution tail comparison sk-T versus PA densities

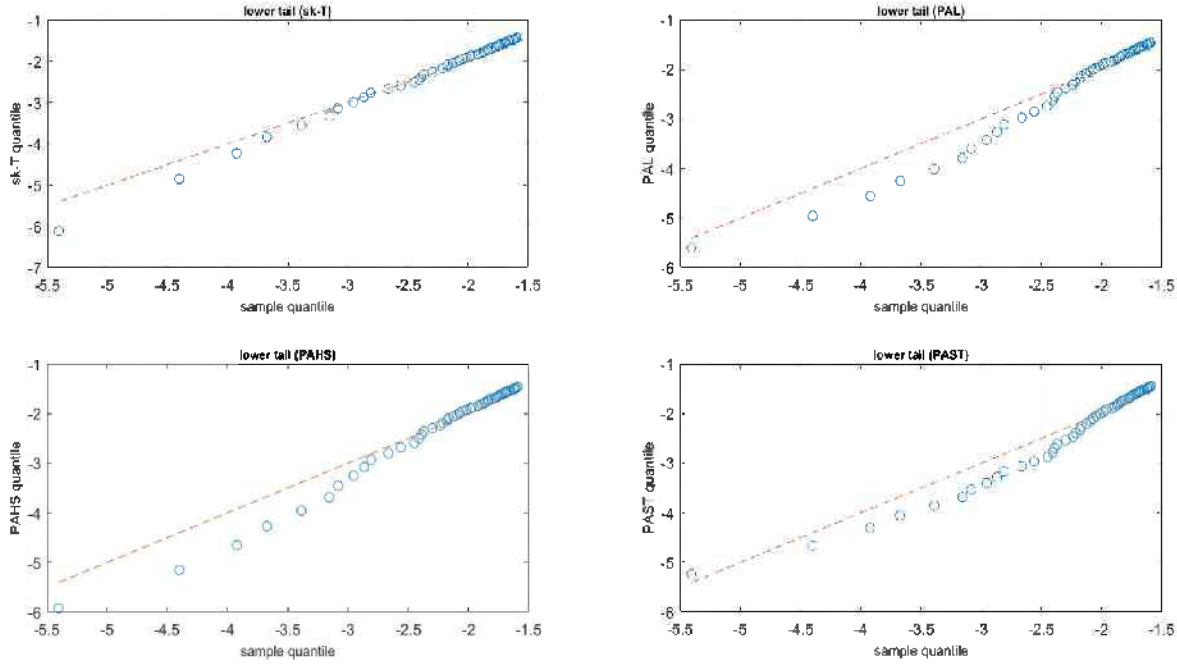


This figure plots theoretical quantiles of sk-T versus PA densities for both tails: lower tail (quantiles from 0.001 to 0.05) and upper tail (quantiles from 0.95 to 0.999). There is a total of fifty equally-spaced quantiles in each tail. All densities have zero mean, unit variance and the same levels of skewness (-0.47) and kurtosis (11).

Figure 5: Distribution tail fit analysis for standardized returns

Series: JAP-US

Panel A: lower tail



Panel B: upper tail

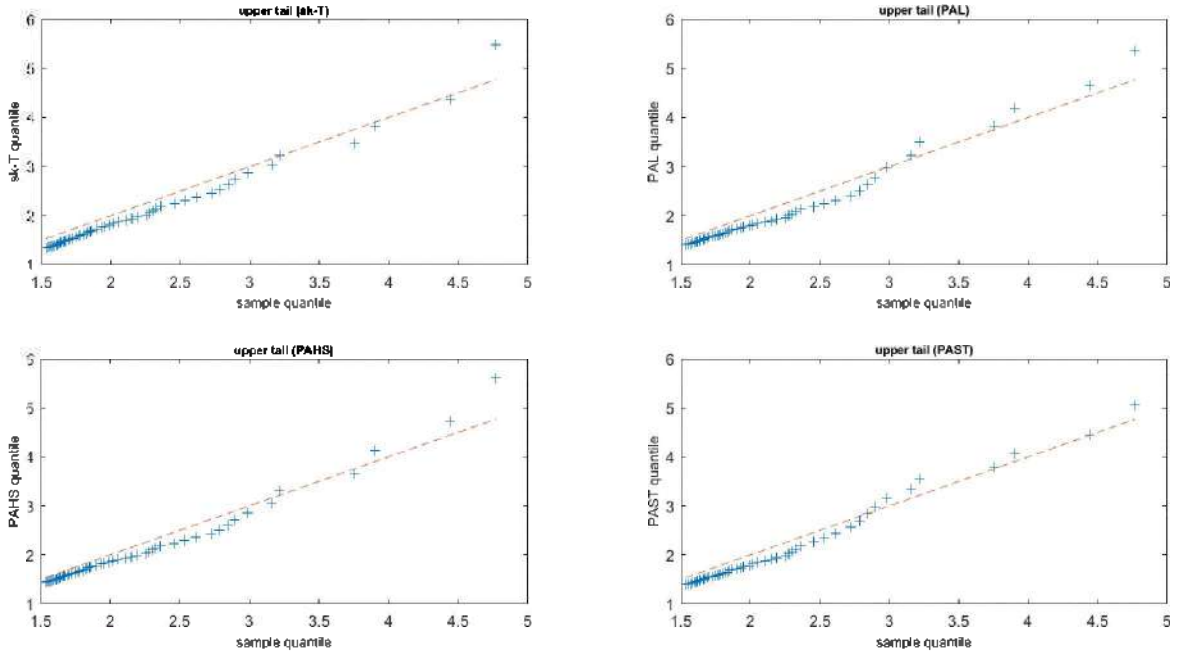
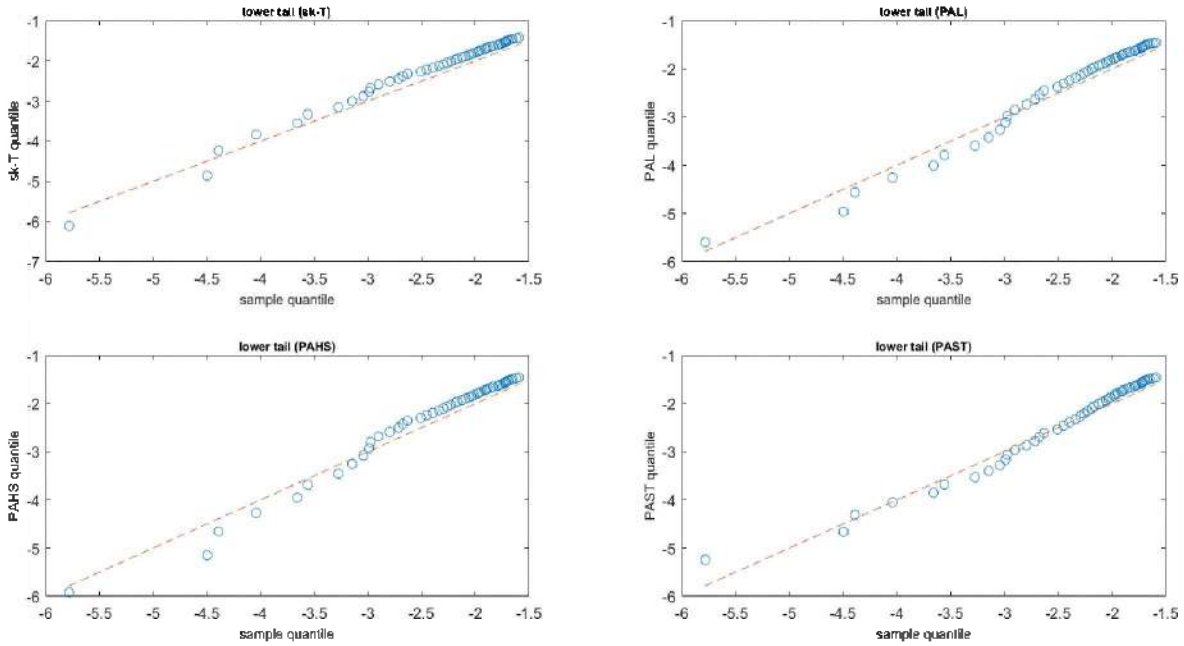


Figure 5 (continued)

Series: GSCITOT

Panel A: lower tail



Panel B: upper tail

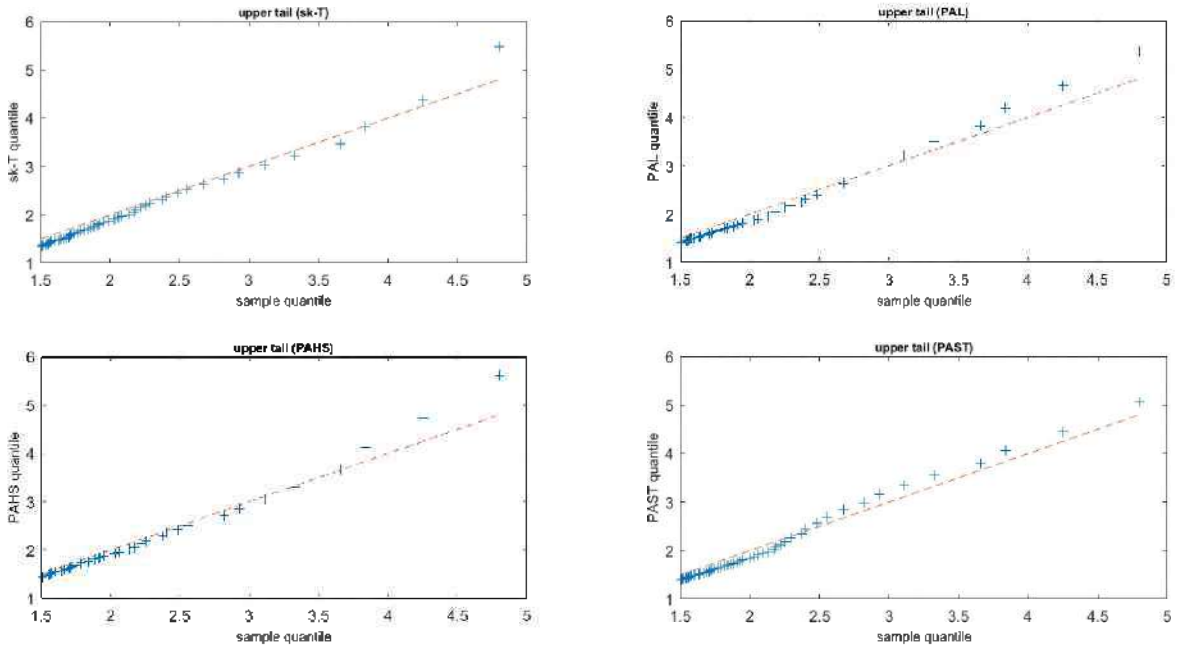
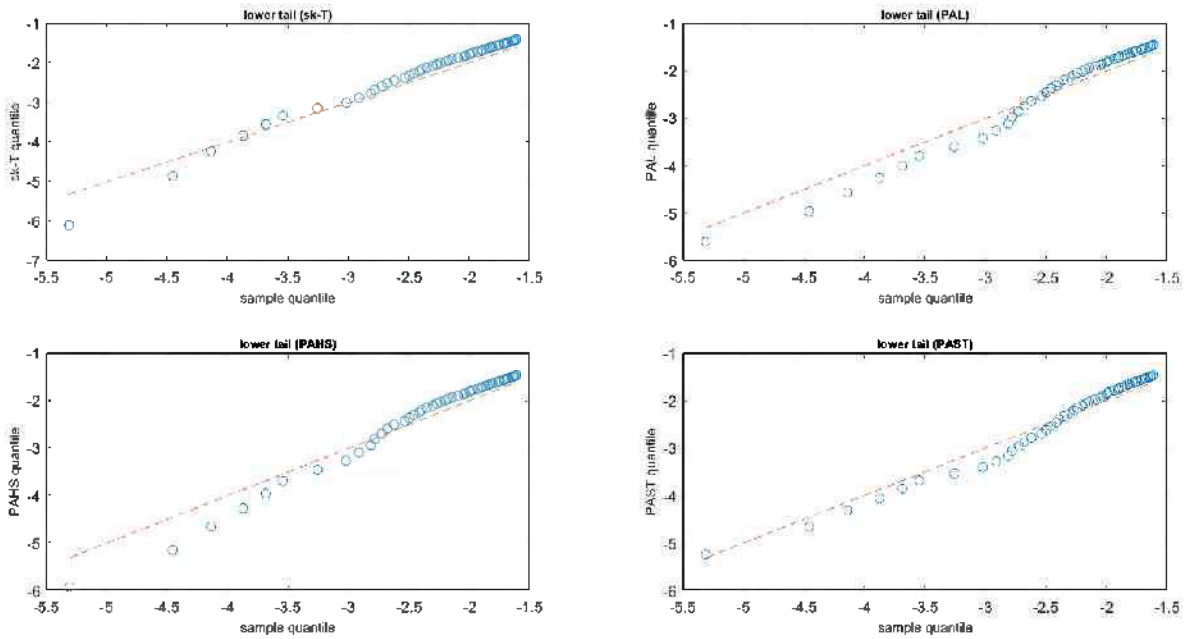


Figure 5 (continued)

Series: GOLD

Panel A: lower tail



Panel B: upper tail

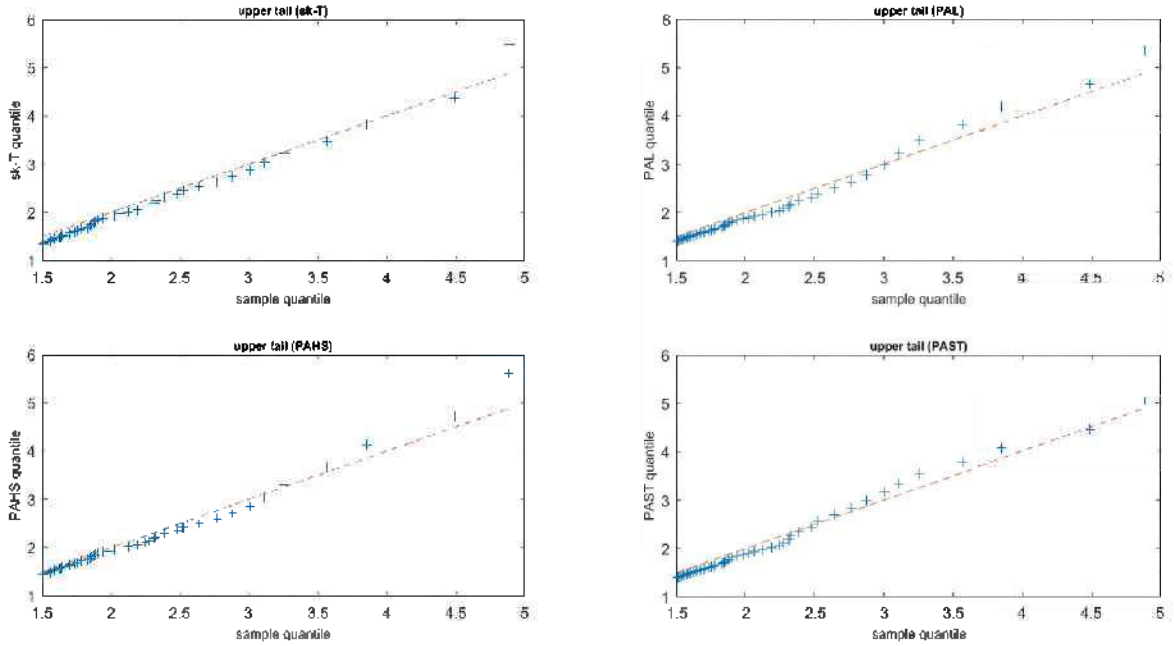
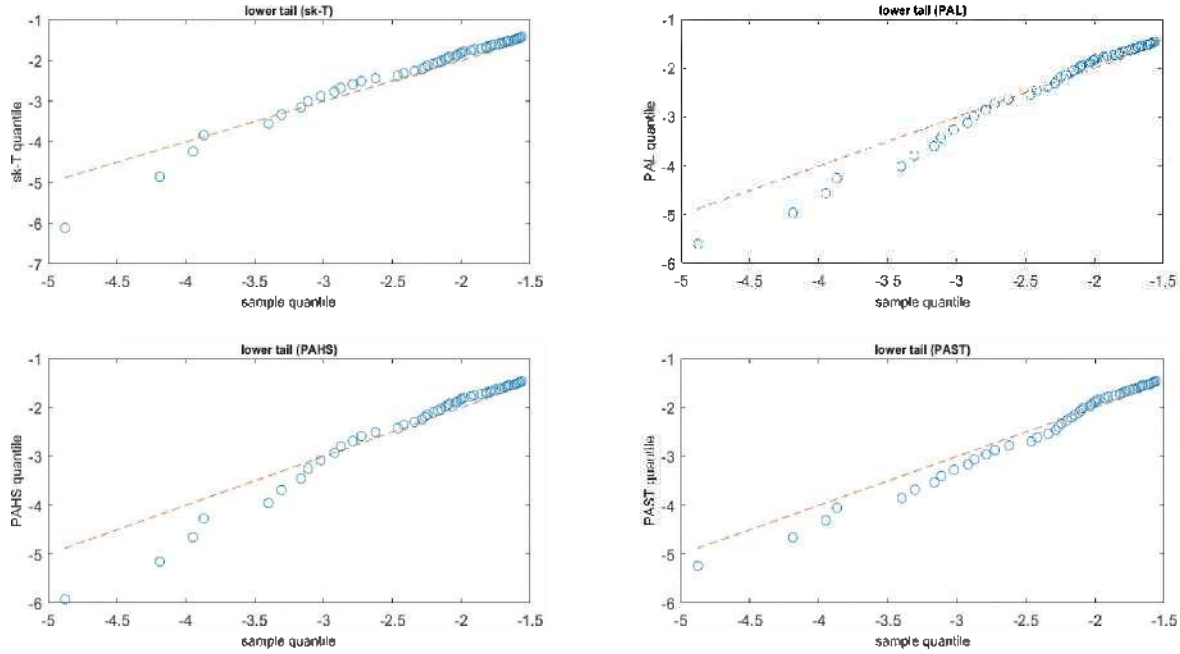


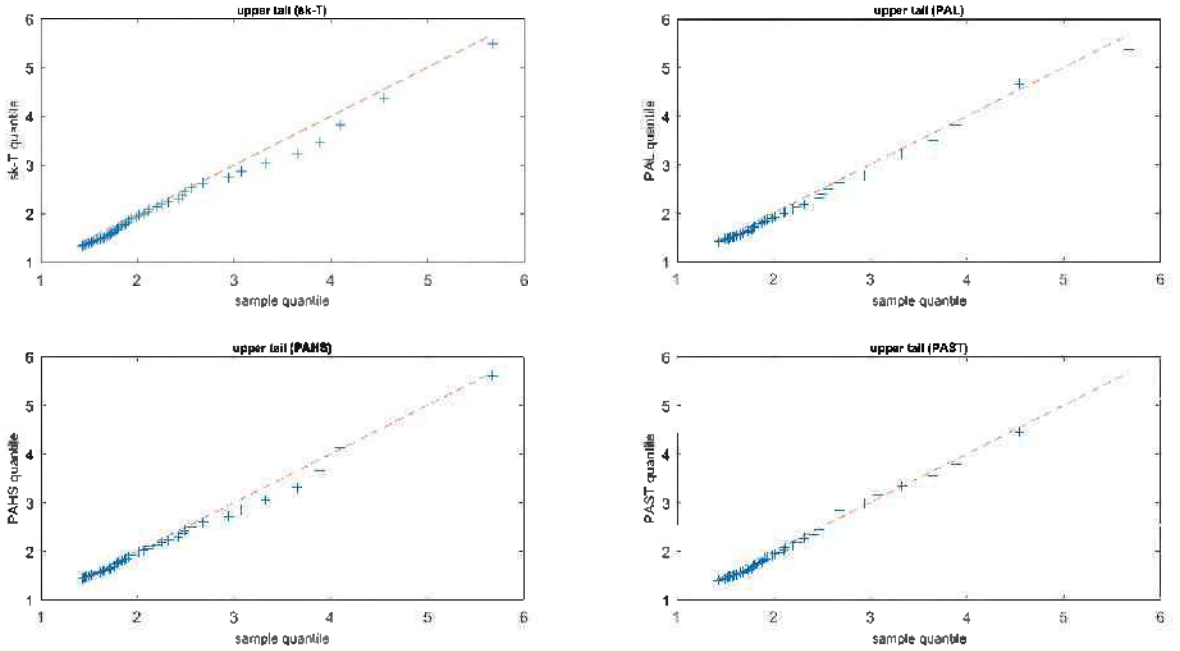
Figure 5 (continued)

Series: BRENT

Panel A: lower tail



Panel B: upper tail



This figure provides sk-T, PAL, PAHS and PAST theoretical quantiles versus sample standardized return quantiles for both distribution tails: lower tail (quantiles from 0.001 to 0.05) and upper tail (quantiles from 0.95 to 0.999). There is a total of 50 equally-spaced quantiles in each tail. Series: JAP-US, GSCITOT, GOLD and BRENT standardized returns ($T = 3,500$ obs.).