The transformed Gram Charlier distribution: Parametric properties and financial risk applications

León, A. and Ñíguez, T.M.

The Transformed Gram Charlier distribution: parametric properties and financial risk applications

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

In this paper we study an extension of the Gram-Charlier (GC) density in Jondeau and Rockinger (2001) which consists of a Gallant and Nychka (1987) transformation to ensure positivity without parameter restrictions. We derive its parametric properties such as unimodality, cumulative distribution, higher-order moments, truncated moments, and the closed-form expressions for the expected shortfall (ES) and lower partial moments. We obtain the analytic k-th order stationarity conditions for the unconditional moments of the TGARCH model under the transformed GC (TGC) density. In an empirical application to asset return series, we estimate the tail index; backtest the density, VaR and ES; and implement a comparative analysis based on Hansen's skewed-t distribution. Finally, we present extensions to time-varying conditional skewness and kurtosis, and a new class of mixture densities based on this TGC distribution.

Keywords: Backtesting; Expected shortfall; Kurtosis; Skewness; Tail index; Unimodality.

JEL classification codes: C2, C5, G1.

*Corresponding author

1 Introduction

Densities based on polynomial expansions (PE) have drawn great attention to model the departures from normality of the empirical return distributions. For instance, in a recent study, Bagnato, Poti and Zoia (2015) present a simple theorem that links the higher-order moments (skewness and kurtosis) of a polynomially expanded parent distribution to those of the resulting distribution. Within this framework, the Gram-Charlier (GC) probability density function (pdf) in Jondeau-Rockinger (2001) (henceforth, JR) is a PE density with the standard Normal as parent pdf. The GC distribution is very tractable theoretically and empirically mainly because its two parameters correspond directly to skewness and excess kurtosis. The following references, without being exhaustive, provide a research line on the development of the GC distribution: Beber and Brandt (2006), Polanski and Stoja (2010), Cheng, Philip, Zhou, Wang and Lo (2011), Ñíguez and Perote (2012), Liu and Luger (2015), Lönnbark (2016), León and Moreno (2017), Zoia, Biffi and Nicolussi (2018) and Del Brio, Mora-Valencia and Perote (2020).[1]

A well-known problem of the GC density function is that it can render negative values. This issue has been mainly dealt with through two approaches. On the one hand, by means of parametric restrictions that ensure the GC pdf has positive probabilities, as in JR (2001). On the other hand, using density function transformations based on the methodology of Gallant and Nychka (1987) (GN henceforth). The latter approach was followed by León, Rubio and Serna (2005) (hereafter LRS) to define a positive GC density function with conditional autoregressive higher-order moments. The LRS model has been proven useful for numerous financial econometric applications; see, for instance, White, Kim and Manganelli (2010), Alizadeh and Gabrielsen (2013), Auer (2015), Gabrielsen, Kirchner, Liu and Zagaglia (2015), Anatolyev and Petukhov (2016), Kräussl, Lehnert and Senulyté (2016), Narayan and Liu (2018) and Wu, Xia and Zhang (2019).

Henceforth, the LRS density will be referred to as the transformed GC (TGC) since it is the result of transforming the GC density in JR (2001) in order to obtain a well-defined density without parameter restrictions. However, unlike the GC density parameters, the TGC ones do lose their direct interpretation as skewness and excess kurtosis. In this paper we study the TGC's parametric properties providing its true higher-order moments and other features of the TGC pdf such as the conditions for unimodality, allowable ranges of skewness and kurtosis, closed-form formulae for (i) the cumulative distribution function (cdf), (ii) one-sided truncated TGC moments, and (iii) asymmetric risk measures such as expected shortfall (ES) and lower partial moments (LPMs).

We illustrate the practical use of this pdf through an application to modeling asset returns. For that purpose, we implement the threshold GARCH (TGARCH) model of Zakoïan (1994) for the conditional volatility. We test the performance of our model through an in-sample analysis, and out-of-sample (OOS) exercises for backtesting VaR and ES as well as for density forecasting. We derive the analytic expressions for the k-th order stationarity conditions for the unconditional moments of the errors under the TGARCH with TGC density for the standardized errors (henceforth, TGC-TGARCH). Indeed, the second and fourth moments will receive special attention and so, the unconditional variance and kurtosis. Besides, we test the finiteness of the unconditional moments through the robust tail-index method of Gabaix and Ibragimov (2011). As a result, we provide stronger evidence respecting the one obtained from the second and fourth order stationarity conditions driven by the TGC-TGARCH parameter estimates. The data we use comprises

[1] Within this literature, a special mention deserves the GC applications to option pricing in Corrado and Su (1996), Corrado (2007), León Mencía and Sentana (2009) (LMS hereafter) and Schlögl (2013).

stock indexes, exchange rates, a commodity and a cryptocurrency. For comparison purposes, we consider both the Normal distribution as the benchmark as well as the symmetric-GC pdf of Zoia et al. (2018), which we refer to as GCK hereafter. Our choice of the latter is based on that, unlike the TGC, the GCK does model directly kurtosis. Hence, in our comparison analyses we can easily isolate the effect of skewness on the performance of the density. Besides, we also compare the TGC's performance with the popular skewed-t distribution of Hansen (1994). Furthermore, a robustness check of the conditional variance under the TGC is done using the nonlinear asymmetric GARCH (NAGARCH) of Engle and Ng (1993) as an alternative to the TGARCH.[2] We evaluate density forecasting performance through p-value discrepancy plots (Davidson and MacKinnon, 1998) together with proper scoring rules in Amisano and Giacomini (2007). VaR and ES forecasting accuracies are tested via the backtesting procedures in Du and Escanciano (2017).

Finally, we highlight two possible extensions left for future research although we show some previous results here. First, we extend the TGC to incorporate time-varying (TV) higher-order moments where the dynamics for the implied TGC parameters are driven by the specification in, among others, Lalancette and Simonato (2017), which accounts for asymmetric responses of conditional skewness and kurtosis to positive and negative shocks. Our in-sample analysis results show that modeling the TGC with TV higher-order moments clearly contributes to improving goodness-of-fit. The empirical evidence from our daily conditional skewness series is also reinforced by using an asymmetry measure beyond the third central moment and specifically, the one based on the return distribution's tails proposed in Jiang, Wu, Zhou and Zhu (2020). Second, we present a mixture of TGC (MTGC hereafter) densities that features higher flexibility than the TGC for capturing large ranges of kurtosis under a rather similar model framework to that in Alexander and Lazar (2006).

The remainder of the paper is structured as follows. Section 2 deals with the GC pdf as a set-up base of our analysis. In Section 3 we characterize the TGC pdf and study its parametric properties. In Section 4 we apply the TGC for modeling the conditional distribution of asset returns under the TGARCH structure, and analyze the power-law tail properties. Section 5 provides an empirical application to asset return series. Section 6 presents two possible extensions of the TGC linked to lines for further research. Section 7 provides a summary of the conclusions. Appendix 1 includes some properties of the Hermite polynomials used throughout the paper. Appendix 2 includes all proofs, and also the k-th order stationarity conditions for the unconditional moments under the NAGARCH model. Appendix 3 presents the estimation results of the NAGARCH under the TGC pdf (henceforth, TGC-NAGARCH). The robust tail-index estimation results for the asset returns are provided in Appendix 4. Finally, Appendix 5 shows a sensitivity analysis based on a simulation procedure that aims to study the theoretical behavior of the implied tail index from Kesten's equation.

2 The GC distribution

The GC pdf is defined according to following the polynomial expansion density:

$$g\left(x,\boldsymbol{\theta}\right)=\phi\left(x\right)\psi(x,\boldsymbol{\theta}), \tag{1}$$

[2] The NAGARCH model is also known as NGARCH, see Christoffersen (2012).

where $x \in \mathbb{R}$, $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ is the parameter vector, $\phi(\cdot)$ is the pdf of the standard Normal distribution and $\psi(\cdot)$ is defined as

$$\psi(x, \boldsymbol{\theta}) = 1 + \frac{\theta_1}{\sqrt{3!}} H_3(x) + \frac{\theta_2}{\sqrt{4!}} H_4(x), \tag{2}$$

such that $H_k(\cdot)$ denote the (normalized) Hermite polynomials in (70) in Appendix 1. The associated cdf, i.e. $G(x, \boldsymbol{\theta}) = \int_{-\infty}^{x} g(u, \boldsymbol{\theta})\, du$, is given by

$$G(x, \boldsymbol{\theta}) = \Phi(x) - \frac{\theta_1}{3\sqrt{2}} H_2(x)\, \phi(x) - \frac{\theta_2}{4\sqrt{3!}} H_3(x)\, \phi(x), \tag{3}$$

where $H_2(x) = \frac{x^2-1}{\sqrt{2}}$, $H_3(x) = \frac{x^3-3x}{\sqrt{3!}}$ and $H_4(x) = \frac{x^4-6x^2+3}{\sqrt{4!}}$. More details about (3) and other properties of the GC distribution can be seen in León and Moreno (2017).

2.1 Higher-order moments

It is verified that the first noncentral moments of x with pdf in (1) are given by $E_g[x] = 0$, $E_g[x^2] = 1$, $E_g[x^3] = \theta_1$ and $E_g[x^4] = \theta_2 + 3$. Thus, x is a standardized random variable (rv) such that θ_1 and θ_2 correspond, respectively, to the skewness, s, and the excess kurtosis, ek, of $g(x, \boldsymbol{\theta})$. We can adopt the following notation: $\theta_1 = s$ and $\theta_2 = ek$. Since $g(\cdot)$ can take negative values for certain values of (s, ek), JR (2001) obtain numerically a restricted space $\boldsymbol{\Gamma}$ for possible values of (s, ek) where the polynomial function of degree four in (2) becomes non-negative for every x, i.e. $\psi(x, \boldsymbol{\theta}) \geq 0$. As a result, the points in $\boldsymbol{\Gamma}$ verify that $0 \leq ek \leq 4$, $|s| \leq 1.0493$ and the range of s in $\boldsymbol{\Gamma}$ depends on the level of ek. The maximum size for skewness is reached for $ek = 2.4508$. From now on, the well-defined GC pdf refers to $g(x, \boldsymbol{\theta})$ subject to $\boldsymbol{\theta} \in \boldsymbol{\Gamma}$. The envelope of $\boldsymbol{\Gamma}$ is exhibited, in continuous-line, in the left panel of Figure 1 (with ek and s in the x-axis and y-axis, respectively).

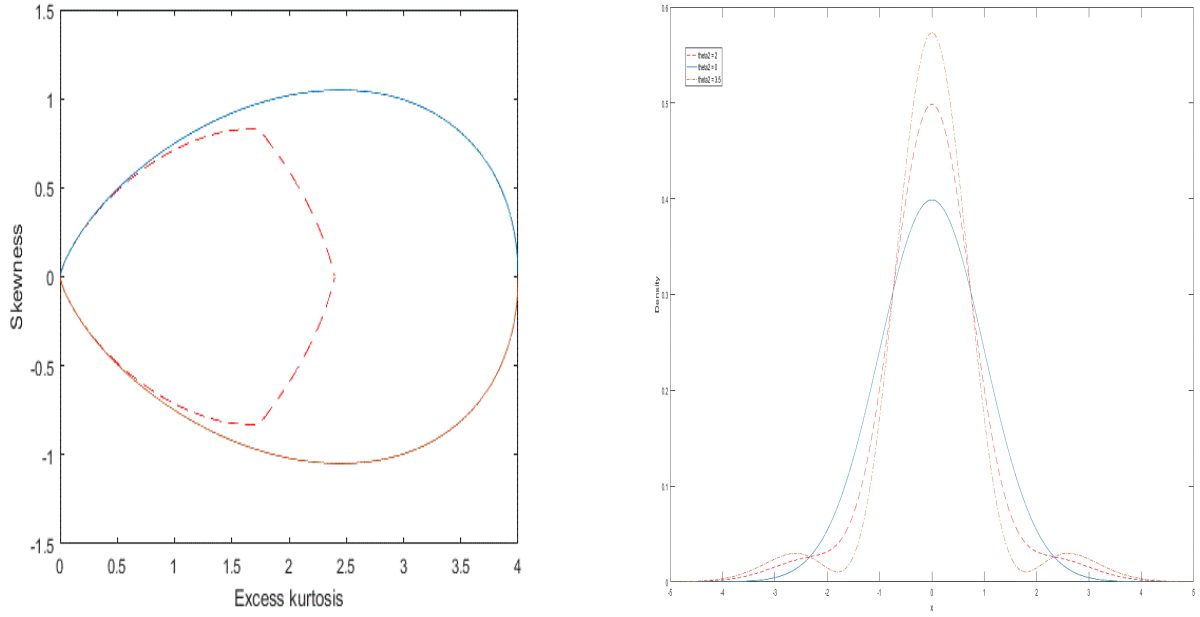


Figure 1: The left panel exhibits the GC positivity frontier (continuous line) and its unimodality frontier (dash line). The right panel plots the GC density for $\theta_1 = 0$ (symmetric distribution) and different values of excess kurtosis, θ_2.

2.2 Unimodality

Figure 1 (left panel) also plots the frontier (dashed-line) of the GC unimodal region, which tally with the curves of Draper and Tierney (1972). The unimodality of $g(\cdot)$ does hold if there is only one real root in the fifth-degree polynomial given by the condition: $x\psi - d\psi/dx = 0$ with $\psi(\cdot)$ in (2). Note that if $g(\cdot)$ is unimodal then $\theta_2 < 2.4$ as shown analytically in Zoia (2010). The right panel of Figure 1 graphs the GC density with $\theta_1 = 0$ (symmetric distribution) and different values of excess kurtosis, θ_2, where unimodality is verified if $\theta_2 = 2$ but not for $\theta_2 = 3.5$. The GC distribution coincides with the standard Normal for the case of $\theta_2 = 0$.

3 The transformed GC distribution

As an alternative to the numerical method implemented in JR (2001) for building the restricted parameter set $\mathbf{\Gamma}$ which ensures the positivity to the pdf in (1), Gallant and Tauchen (1989) suggested to square the polynomial component $\psi(\cdot, \boldsymbol{\theta})$ in (2). As a result, we can obtain a new pdf $q(\cdot)$, which we call the transformed GC (TGC), and given by

$$q(x, \boldsymbol{\theta}) = \lambda \phi(x) \psi^2(x, \boldsymbol{\theta}), \tag{4}$$

where the parameter λ verifies that the pdf in (4) is well-defined and hence, the integral of $q(\cdot)$ must be equal to one. The inverse of λ is given by the expression: $1/\lambda = 1 + \gamma_1^2 + \gamma_2^2$ where $\gamma_1 = \theta_1/\sqrt{3!}$ and $\gamma_2 = \theta_2/\sqrt{4!}$. Note that, by transforming $g(\cdot)$ into $q(\cdot)$ the parameters in $q(\cdot)$ are not restricted now. However, by doing so and unlike the GC pdf, they cannot be interpreted as higher moments of the new density. Both skewness and kurtosis under (4) are indeed non-linear functions of θ_1 and θ_2. If we expand the square of $\psi(x)$, we can rewrite (4) as

$$q(x, \boldsymbol{\theta}) = \lambda \phi(x) \left[1 + 2\gamma_1 H_3(x) + 2\gamma_2 H_4(x) + 2\gamma_1\gamma_2 H_3(x) H_4(x) + \gamma_1^2 H_3^2(x) + \gamma_2^2 H_4^2(x)\right]. \tag{5}$$

Note that (4) is nested in a more general pdf that belongs to the SNP class introduced by GN (1987) and, also, by LMS (2009) who studied its parametric properties. Thus,

$$p_n(x, \boldsymbol{\nu}) = \frac{\phi(x)}{\boldsymbol{\nu}'\boldsymbol{\nu}} \left(\sum_{k=0}^{n} \nu_k H_k(x)\right)^2, \tag{6}$$

where $\boldsymbol{\nu} = (\nu_0, \nu_1, \ldots, \nu_n)' \in \mathbb{R}^{n+1}$ and $\nu_0 = 1$ to solve the scale indeterminacy in (6). Definitively, the pdf in (6) directly nests (4) when $n = 4$, $\nu_1 = 0$, $\nu_2 = 0$ and $\nu_k = \gamma_k$ for $k = 3, 4$. In short $q(x, \boldsymbol{\theta})$ is a restricted model of $p_4(x, \boldsymbol{\nu})$ when $\nu_1 = \nu_2 = 0$.

Proposition 1. *Consider the pdf $q(\cdot)$ in (5), then*

$$\begin{aligned} \Psi_k(x) &= \lambda^{-1} E_q\left[H_k(u)\mathcal{I}(u \leq x)\right] = \lambda^{-1}\int_{-\infty}^{x} H_k(u) q(u, \boldsymbol{\theta})\, du \\ &= \Gamma_{k00}(x) + 2\gamma_1\Gamma_{k30}(x) + 2\gamma_2\Gamma_{k40}(x) + 2\gamma_1\gamma_2\Gamma_{k34}(x) + \gamma_1^2\Gamma_{k33}(x) + \gamma_2^2\Gamma_{k44}(x), \end{aligned} \tag{7}$$

where $k \in \mathbb{N}$ and $\Gamma_{kij}(x) = E_\phi\left[H_k(u) H_i(u) H_j(u)\mathcal{I}(u \leq x)\right]$ such that $\mathcal{I}(A) = 1 \Leftrightarrow A$ is verified (otherwise, $\mathcal{I}(A) = 0$). The general expression of $\Gamma_{kij}(\cdot)$ can be rewritten as $\Gamma_{kij}(x) = \sum_{l=0}^{n} \omega_{n,l} B_l(x)$, such that $n = k + i + j$, $\omega_{n,l}$ are coefficients, and $B_l(x) = \int_{-\infty}^{x} u^l \phi(u) du$ is given in (73), see Appendix 2.

Proof. The equation (7) is obtained straightforwardly. ■

The cdf for the TGC distribution is obtained as

$$Q(x,\boldsymbol{\theta}) = \int_{-\infty}^{x} q(u,\boldsymbol{\theta})\,du = \lambda\Psi_0(x), \tag{8}$$

where $\Psi_0(\cdot)$ is given in (7) for $k=0$. Note that $Q(x,\boldsymbol{\theta}) = \Gamma_{000}(x) = \Phi(x)$ is the standard Normal cdf. The expressions of $\Gamma_{0ij}(\cdot)$ in $\Psi_0(\cdot)$ are in Appendix 2.

3.1 Higher-order moments

Proposition 2. *The first four noncentral moments of the random variable x with* (5) *as pdf are given by*

$$\begin{aligned} E_q[x] &= 4\lambda\gamma_1\gamma_2, & E_q\left[x^3\right] &= 2\sqrt{6}\lambda\gamma_1 + 48\lambda\gamma_1\gamma_2, \\ E_q\left[x^2\right] &= 6\lambda\gamma_1^2 + 8\lambda\gamma_2^2 + 1, & E_q\left[x^4\right] &= 4\sqrt{6}\lambda\gamma_2 + 72\lambda\gamma_1^2 + 120\lambda\gamma_2^2 + 3. \end{aligned} \tag{9}$$

Proof. See Appendix 2. ■

Let z denote the standardized rv of x, then $z = a(\boldsymbol{\theta}) + b(\boldsymbol{\theta})x$ where $a = -bE_q[x]$, $b = 1/\sigma_x$ and $\sigma_x^2 = E_q\left[x^2\right] - E_q[x]^2$. Hence, the pdf of z is obtained as $\frac{1}{b}q\left(\frac{z-a}{b}\right)$. Both skewness and kurtosis of z are given by

$$s_z = E_q\left[(a+bx)^3\right] = a^3 + 3a^2bE_q[x] + 3ab^2E_q\left[x^2\right] + b^3E_q\left[x^3\right], \tag{10}$$

$$k_z = E_q\left[(a+bx)^4\right] = a^4 + 4a^3bE_q[x] + 6a^2b^2E_q\left[x^2\right] + 4ab^3E_q\left[x^3\right] + b^4E_q\left[x^4\right]. \tag{11}$$

Figure 2 exhibits the skewness and excess kurtosis region (shaded area) for the TGC distribution given the above equations (10) and (11). Note that it contains part of the GC envelope. It is also displayed in dashed-line the skewness-excess kurtosis boundary (for a standardized distribution) ensuring the existence of a density, i.e. $s < \pm\sqrt{ek+2}$. If we consider a grid for $\theta_j \in [-20,20]$ with length of 0.01 where $j=1,2$, then the TGC verifies that $-1.4536 \leq ek_z \leq 2.7208$ and $|s_z| \leq 1.2224$. Hence, it does allow thinner tails than those of the Normal distribution. The maximum size of s_z is reached for $ek_z = 1.0643$. Both maximum and minimum values of ek_z are obtained for $s_z = 0$. It can be seen that levels of ek_z larger than 2.7208 are not captured under the TGC and so, it is less flexible than the GC for more leptokurtic distributions.

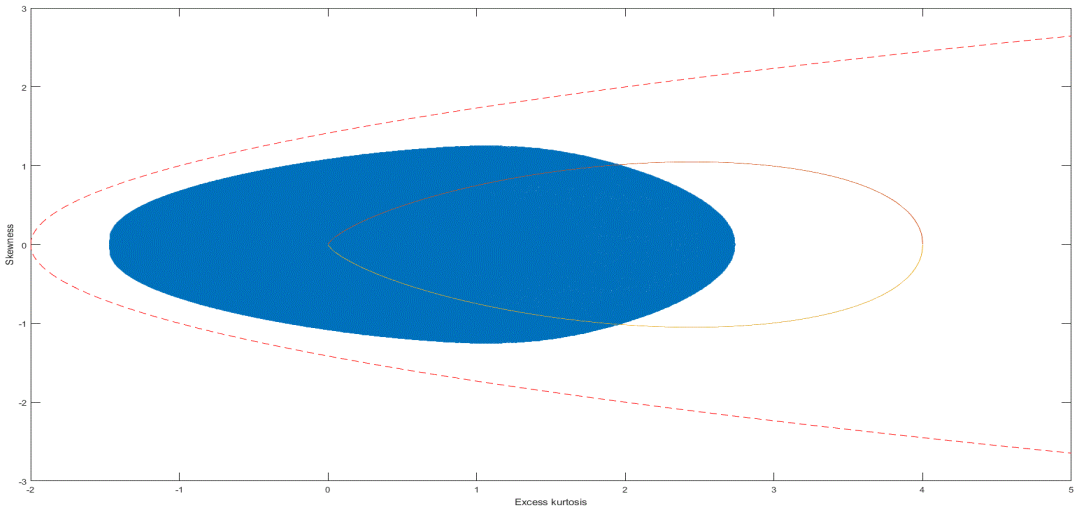


Figure 2: Allowable skewness and excess kurtosis region for TGC pdf (shaded area). The dash-line represents the skewness-excess kurtosis boundary. The GC envelope is represented by the continuous line.

Figure 3 depicts the skewness function (10). The left panel displays the range of s_z as θ_1 varies between -5 and 5 given some selected values of θ_2, i.e. $s_z\left(\theta_1, \overline{\theta}_2\right)$ where $\overline{\theta}_2 \in \{-3, 0, 3\}$. Note that the graph of $s_z\left(\theta_1, \overline{\theta}_2\right)$ behaves like an odd function with respect to θ_1: $s_z\left(-\theta_1, \overline{\theta}_2\right) = -s_z\left(\theta_1, \overline{\theta}_2\right)$. The sign of s_z coincides with that of θ_1 for $\overline{\theta}_2 = 0$ as exhibited in the curve $s_z\left(\theta_1, 0\right)$. The maximum size of s_z is also obtained for $\overline{\theta}_2 = 0$. The right panel shows the dynamics of s_z as θ_2 varies while θ_1 is fixed, i.e. $s_z\left(\overline{\theta}_1, \theta_2\right)$ with selected values for θ_1 just the same as those for θ_2 in the left panel. It is verified that $s_z\left(0, \theta_2\right) = 0$ and $s_z\left(-\overline{\theta}_1, \theta_2\right) = -s_z\left(\overline{\theta}_1, \theta_2\right)$ when $\overline{\theta}_1 \neq 0$ (symmetry respecting the x-axis, denoted as θ_2). We can see that s_z decreases (increases) if θ_2 increases for $\overline{\theta}_1 = -3$ $\left(\overline{\theta}_1 = 3\right)$.

Figure 4 is constructed in the same way as Figure 3 but now it exhibits the excess kurtosis function, ek_z, with k_z defined in (11). The left and right panels display $ek_z\left(\theta_1, \overline{\theta}_2\right)$ and $ek_z\left(\overline{\theta}_1, \theta_2\right)$, respectively. Note that in the left panel, the graph $ek_z\left(\theta_1, \overline{\theta}_2\right)$ behaves like an even function with respect to θ_1: $ek_z\left(-\theta_1, \overline{\theta}_2\right) = ek_z\left(\theta_1, \overline{\theta}_2\right)$. Hence, the sign of θ_1 does not influence the behavior of ek. As a result, we can see in the right panel that $ek_z\left(\overline{\theta}_1, \theta_2\right) = ek_z\left(-\overline{\theta}_1, \theta_2\right)$ when $\overline{\theta}_1 \neq 0$. Higher values of ek are obtained (in most cases) for $\overline{\theta}_1 = 0$.

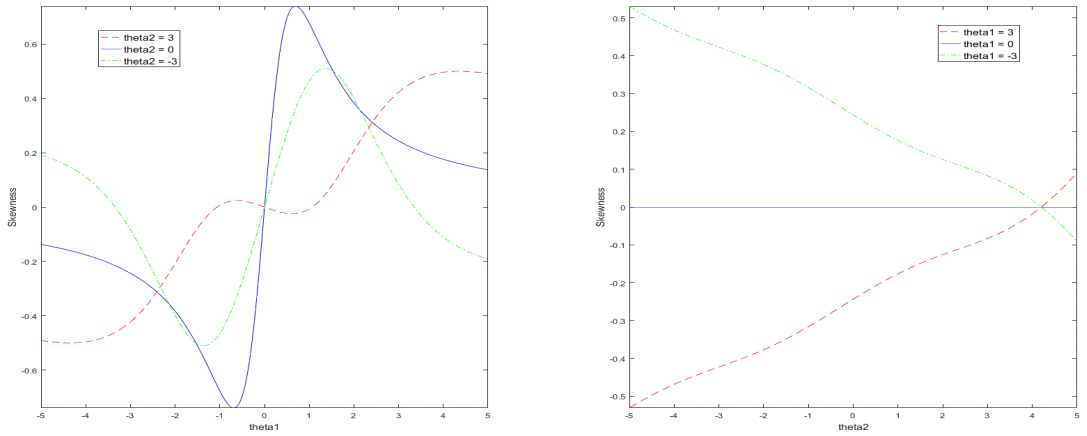


Figure 3: The left panel displays the skewness function $s_z\left(\theta_1, \overline{\theta}_2\right)$ for θ_1 given $\overline{\theta}_2 \in \{-3, 0, 3\}$. The right panel shows the range of skewness function $s_z\left(\overline{\theta}_1, \theta_2\right)$ for θ_2 given $\overline{\theta}_1 \in \{-3, 0, 3\}$.

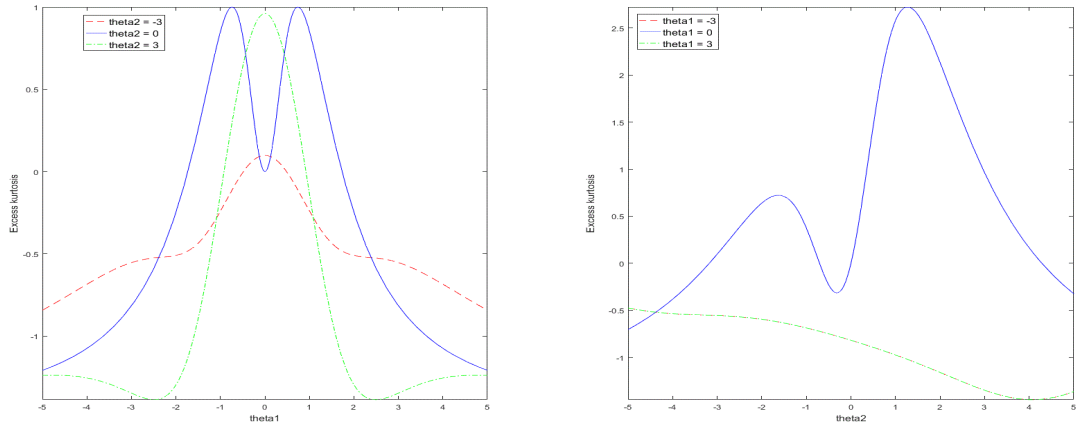


Figure 4: The left and right panels display excess-kurtosis functions $ek_z\left(\theta_1, \overline{\theta}_2\right)$ and $ek_z\left(\overline{\theta}_1, \theta_2\right)$, for $\overline{\theta}_1$, $\overline{\theta}_2 \in \{-3, 0, 3\}$, respectively.

3.2 Unimodality

The unimodality of $q(\cdot)$ in (5) does hold if there is only one real root in the ninth-degree polynomial given by the condition: $2\psi d\psi/dx - x\psi^2 = 0$ with $\psi(\cdot)$ in (2). The left panel in Figure 5 contains the TGC unimodal region such that the unimodality is verified for $0 \leq ek < 2.7$. Note that the unimodality property leads to a slightly upper bound for ek under TGC than GC as can be exhibited when also plotting the GC frontier under unimodality. The right panel in Figure 5 exhibits the values of θ_1 and θ_2 such that the unimodality is verified.

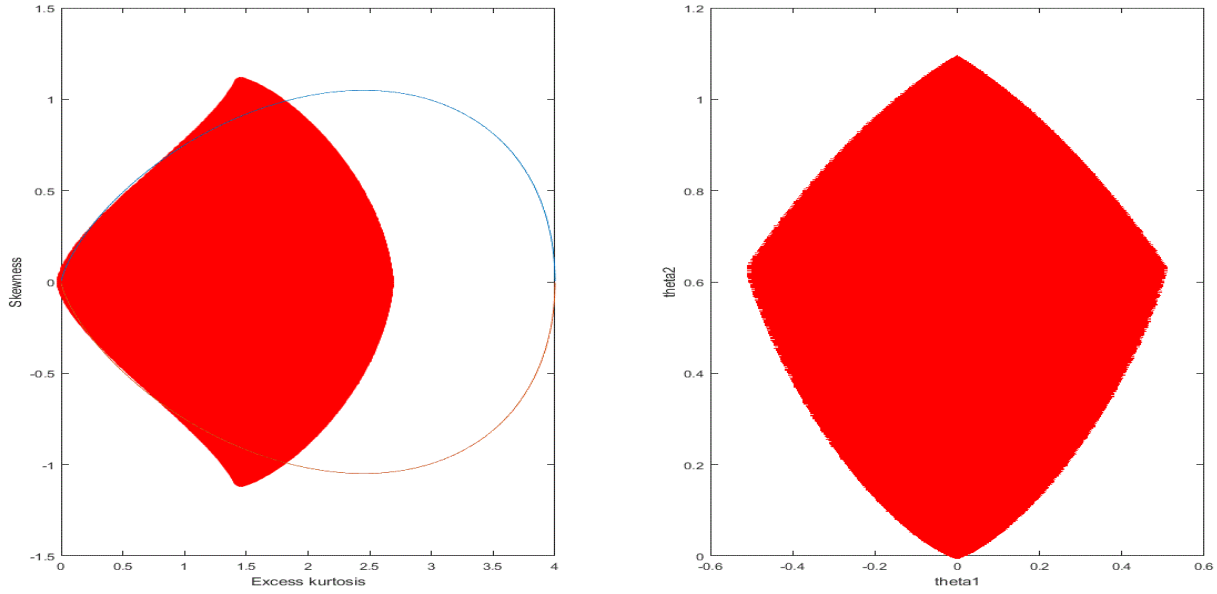


Figure 5: The left panel contains the TGC unimodal region (shaded area) in terms of skewness-excess kurtosis. The right panel exhibits the values of θ_1 and θ_2 such that unimodality holds. The dash line is the usual GC envelope.

Note that (10) and (11) are non-linear functions of θ_1 and θ_2. We are interested in studying the sensitivity of these higher moments with respect to both parameters, and for the sake of simplicity we consider the unimodality region. For that purpose, we simplify the non-linear relationship by adjusting a multivariate polynomial curve fitting to each higher-order moment series. Specifically, following Amédée-Manesme, Barthélémy and Maillard (2019), we implement a quadratic response surface model, i.e.

$$y = \beta_0 + \beta_1\theta_1 + \beta_2\theta_2 + \beta_3\theta_1^2 + \beta_4\theta_2^2 + \beta_5\theta_1\theta_2 + \varepsilon, \tag{12}$$

where y denotes either s_z or k_z,[3] and ε is a random variable with $E(\varepsilon) = 0$ and $V(\varepsilon) = \sigma_\varepsilon^2$. It is verified that the R-squared is very high in both least squares regressions and all coefficients are significant. These

[3] Because of symmetry in Figure 5 (left panel) respecting the x-axis, we run equation (12) with $s_z > 0$ and k_z as dependent variables. Similar conclusions are obtained for $s_z < 0$.

results are available upon request. Next, we obtain the elasticity measures defined as $E_{y,\theta_i} = \frac{\partial y}{\partial \theta_i}\frac{\theta_i}{y}$, where $\frac{\partial y}{\partial \theta_i} = \beta_i + 2\beta_{i+2}\theta_i + \beta_5\theta_j$ for $i = 1, 2$ and $i \neq j$. The main results are the following. First, $E_{s_z,\theta_1} > 0$ and $E_{k_z,\theta_2} > 0$ for all points (θ_1, θ_2) that belong to the unimodal region. Second, $|E_{s_z,\theta_1}| > |E_{s_z,\theta_2}|$ and $|E_{k_z,\theta_2}| > |E_{k_z,\theta_1}|$ in most cases. Finally, although both θ_1 and θ_2 influence s_z and k_z, the former seems to be more relevant for skewness while the latter is so for kurtosis.

3.3 Asymmetric risk measures

We obtain the closed-form expressions under the TGC distribution for both ES and LPM.

3.3.1 VaR and Expected shortfall

Let $x_\alpha = Q^{-1}(\alpha)$ denote the α-quantile (or VaR), i.e. $Q^{-1}(\alpha) = \inf\{x \,|Q(x, \boldsymbol{\theta}) \geq \alpha\}$. The ES of the random variable x with pdf $q(\cdot)$ in (5) is given by

$$ES_q(\alpha) = E_q[x \,|x \leq x_\alpha] = \alpha^{-1}\lambda\Psi_1(x_\alpha), \tag{13}$$

where $\Psi_1(\cdot)$ is obtained in (7) for $k = 1$, and the expressions of $\Gamma_{1ij}(\cdot)$ in $\Psi_1(\cdot)$ are in Appendix 2. Finally, the ES under the GC distribution with pdf $g(\cdot)$ in (1) and $x_\alpha = G^{-1}(\alpha)$ as the α-quantile with cdf $G(\cdot)$ in (3), is easily obtained as

$$ES_g(\alpha) = E_g[x \,|x \leq x_\alpha] = \frac{1}{\alpha}\Gamma_{100}(x_\alpha) + \frac{\gamma_1}{\alpha}\Gamma_{130}(x_\alpha) + \frac{\gamma_2}{\alpha}\Gamma_{140}(x_\alpha). \tag{14}$$

3.3.2 Lower partial moments

The LPMs, see Fishburn (1977), of order m for x with pdf in (5) and threshold of τ is defined as

$$LPM_q(\tau, m) = \int_{-\infty}^{\tau} (\tau - x)^m q(x, \boldsymbol{\theta})\, dx. \tag{15}$$

The LPM of order 1 is easily obtained as

$$LPM_q(\tau, 1) = \tau Q(\tau, \boldsymbol{\theta}) - Q(\tau, \boldsymbol{\theta})\, ES_q(Q(\tau, \boldsymbol{\theta})), \tag{16}$$

with $ES_q(Q(\tau, \boldsymbol{\theta})) = E_q[x \,|x \leq \tau]$ given in (13). The LPM of order 2 is

$$LPM_q(\tau, 2) = \tau^2 Q(\tau, \boldsymbol{\theta}) - 2\tau Q(\tau, \boldsymbol{\theta})\, ES_q(Q(\tau, \boldsymbol{\theta})) + E_q\left[x^2 \mathcal{I}(x \leq \tau)\right], \tag{17}$$

where $E_q\left[x^2\mathcal{I}(x \leq \tau)\right] = \sqrt{2}\lambda\Psi_2(\tau) + Q(\tau, \boldsymbol{\theta})$, such that $\Psi_2(\cdot)$ is obtained in (7) for $k = 2$, and the expressions of $\Gamma_{2ij}(\cdot)$ in $\Psi_2(\cdot)$ are in Appendix 2. Note that, $E_q\left[x^k\mathcal{I}(x \leq \tau)\right]$ where $k \in \mathbb{N} \setminus \{0\}$ can easily be rewritten in terms of the one-sided k-th truncated moment, i.e. $E_q\left[x^k\mathcal{I}(x \leq \tau)\right] = Q(\tau, \boldsymbol{\theta})\, E_q\left[x^k \,|x \leq \tau\right]$.

3.4 Alternative positive transformations

The SNP density in (6) nests some other GC expansion transformations used in the literature so as to ensure positivity. First, we consider the case of squaring the terms of equation (2). This approach can be seen in Ñíguez and Perote (2012) with pdf defined as

$$\widetilde{q}\left(x,\boldsymbol{\nu}\right)=\lambda\phi\left(x\right)\left[1+\nu_{1}^{2}H_{3}^{2}(x)+\nu_{2}^{2}H_{4}^{2}(x)\right], \tag{18}$$

where $\boldsymbol{\nu}=\left(\nu_{1},\nu_{2}\right)'\in\mathbb{R}^{2}$ and $1/\lambda=1+\nu_{1}^{2}+\nu_{2}^{2}$. The non-central moments are easily obtained in the following result:

Corollary 1. *The first four noncentral moments of the random variable x with* (18) *as pdf are given by*

$$\begin{aligned} &E_{\widetilde{q}}\left[x\right]=0, && E_{\widetilde{q}}\left[x^{3}\right]=0,\\ &E_{\widetilde{q}}\left[x^{2}\right]=6\lambda\nu_{1}^{2}+8\lambda\nu_{2}^{2}+1, && E_{\widetilde{q}}\left[x^{4}\right]=72\lambda\nu_{1}^{2}+120\lambda\nu_{2}^{2}+3. \end{aligned} \tag{19}$$

Proof. It is obtained straightforwardly given some results of Proposition 2. ■

Let z denote again the standardized rv of x with (18) as pdf, then $z=bx$ where $b=1/\sqrt{E_{\widetilde{q}}\left[x^{2}\right]}$. Hence, the skewness of z is zero and the kurtosis of z is given by

$$k_{z}=\frac{3\left(24\lambda\nu_{1}^{2}+40\lambda\nu_{2}^{2}+1\right)}{\left(6\lambda\nu_{1}^{2}+8\lambda\nu_{2}^{2}+1\right)^{2}}. \tag{20}$$

It can be shown that the maximum and minimum levels of ek_z are about 1.69 and -1.47, respectively. Note that $ek_z=1.69$ is much lower than the maximum value of $ek_z=2.72$ under the TGC distribution with pdf in (5).

Second, if one only aims to capture higher excess kurtosis levels under these kind of positive transformations, then an easy approach can be the following restricted SNP density:

$$q_{n}\left(x,\nu_{n}\right)=\lambda\phi\left(x\right)\left[1+\nu_{n}^{2}H_{n}^{2}(x)\right], \tag{21}$$

where $\nu_{n}\in\mathbb{R}$ and $1/\lambda=1+\nu_{n}^{2}$. Note that (21) can be obtained by eliminating all squared hermite polynomials $H_{k}^{2}(x)$ such that $0<k<n$ and all products $H_{i}(x)H_{j}(x)$ with $i\neq j$ from the SNP pdf in (6). Note that all odd moments related to (21) are equal to zero.

Corollary 2. *Let z* be *the standardized rv of x with* (21) *as pdf, i.e. $z=bx$* where $b=1/\sqrt{E_{q_{n}}\left[x^{2}\right]}$, *then the kurtosis of z is given by*

$$k_{z}\left(n\right)=\frac{2\sqrt{6}\lambda\nu_{n}^{2}\left(A_{4nn}+\sqrt{3}A_{2nn}\right)+3}{\left(\sqrt{2}\lambda\nu_{n}^{2}A_{2nn}+1\right)^{2}}, \tag{22}$$

where $A_{jnn}=E_{\phi}\left[H_{j}\left(x\right)H_{n}^{2}\left(x\right)\right]$.

Proof. It is obtained straightforwardly from some results of Proposition 2. ■

It can be shown that the excess kurtosis in (22) does increase with n such that ek_z ranges from 0 ($n = 1$) to about 6.1 ($n = 10$).

4 Model for returns

We assume the asset return process r_t is defined as $r_t = \mu_t + \varepsilon_t$ with $\varepsilon_t = \sigma_t z_t$, where μ_t and σ_t^2 denote the conditional mean and variance of r_t given by $\mu_t = E\left[r_t \left| I_{t-1} \right.\right]$ and $\sigma_t^2 = E\left[\left(r_t - \mu_t\right)^2 \left| I_{t-1} \right.\right]$ such that I_{t-1} is the information set available at $t-1$ and z_t are the innovations with zero mean, unit variance and D_t as the distribution with TV parameter set, i.e. $z_t \sim D_t\left(0,1\right)$. Note that D_t nests the simple case of constant parameters across time of the distribution of z_t, i.e. $z_t \sim i.i.d.\ D\left(0,1\right)$. We adopt alternative distributions for z_t. Respecting the conditional variance, we start considering the TGARCH of Zakoïan (1994), which models directly the volatility σ_t instead of σ_t^2 and provides for the leverage effect. In short, we model the return series $\{r_t\}$ as

$$r_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t = \sigma_t z_t, \qquad z_t \sim D_t\left(0,1\right), \tag{23}$$

$$\sigma_t = \alpha_0 + \beta\sigma_{t-1} + \alpha_1^+ \varepsilon_{t-1}^+ - \alpha_1^- \varepsilon_{t-1}^-, \tag{24}$$

such that $\alpha_0 > 0$, $\beta \geq 0$, $\alpha_1^+ \geq 0$ and $\alpha_1^- \geq 0$. We use the notation: $\varepsilon_t^+ = \max\left(\varepsilon_t, 0\right)$, $\varepsilon_t^- = \min\left(\varepsilon_t, 0\right)$.

4.1 Moment conditions of TGARCH under TGC density

We can rewrite (24) as

$$\sigma_t = \alpha_0 + c_{t-1}\sigma_{t-1}, \tag{25}$$

where

$$c_t = \beta + \alpha_1^+ z_t^+ - \alpha_1^- z_t^-. \tag{26}$$

Following He and Teräsvirta (1999), and Franq and Zakoian (2010) (FZ hereafter) for the existence of the k-th moment of ε_t in (23), it is necessary that $E\left(\left|z_t\right|^k\right) < \infty$ and also the k-th order stationarity condition of σ_t, i.e. $E\left(\sigma_t^k\right) < \infty$ with σ_t in (25) if and only if $E\left(c_t^k\right) < 1$. The following result provides recursively the expressions of $E\left(c_t^k\right)$, and henceforth denoted as ϖ_k.

Proposition 3. *Let* $\varpi_k = E\left(c_t^k\right)$ *with* c_t *in* (26) *for the TGARCH model* (25) *with* $z_t \sim i.i.d.\ D\left(0,1\right)$, *then for* $k \geq 2$:

$$\varpi_k = \left(\alpha_1^+\right)^k E\left(z_t^k\right) + \left[\left(-1\right)^k \left(\alpha_1^-\right)^k - \left(\alpha_1^+\right)^k\right] E\left[\left(z_t^-\right)^k\right] - \sum_{j=1}^{k} \binom{k}{j} \left(-\beta\right)^j \varpi_{k-j}, \tag{27}$$

where

$$\varpi_1 = \beta - \left(\alpha_1^- + \alpha_1^+\right) E\left(z_t^-\right). \tag{28}$$

Proof. Consider $E\left(z_t^k\right) = E\left[\left(z_t^+\right)^k\right] + E\left[\left(z_t^-\right)^k\right]$ and $E\left[\left(z_t^+\right)^i\left(z_t^-\right)^j\right] = 0 \;\; \forall i, j > 0$, then we easily obtain (27). ■

Corollary 3. *The TGARCH unconditional variance of ε_t in (23), denoted as $\sigma_\varepsilon^2 = E\left(\varepsilon_t^2\right)$, with $\varpi_2 < 1$ and $z_t \sim$ i.i.d. $D(0,1)$ is given by*

$$\sigma_\varepsilon^2 = E\left(\sigma_t^2\right) = \frac{\alpha_0^2\left(1+\varpi_1\right)}{\left(1-\varpi_1\right)\left(1-\varpi_2\right)}, \tag{29}$$

where ϖ_1 is the equation (28) and

$$\varpi_2 = \beta^2 + \left(\alpha_1^+\right)^2 + \left[\left(\alpha_1^-\right)^2 - \left(\alpha_1^+\right)^2\right] E\left[\left(z_t^-\right)^2\right] - 2\beta\left(\alpha_1^- + \alpha_1^+\right) E\left(z_t^-\right). \tag{30}$$

Proof. Following FZ (2010) it is easily obtained (29), and (30) by using (27).

Proposition 4. *Let $z_t \sim$ i.i.d. $TGC(0,1,\boldsymbol{\theta})$ in (23) represents the standardized TGC distribution, i.e. $z_t = a + bx_t \sim$ i.i.d. $D(0,1)$ such that $x_t \sim TGC(\boldsymbol{\theta})$ with pdf in (5). Hence, the general expression of $E\left[\left(z_t^-\right)^k\right]$ is obtained as*

$$\begin{aligned} E\left[\left(z_t^-\right)^k\right] &= \int_{-\infty}^{-a/b}(a+bx_t)^k\, q(x_t)\, dx_t \\ &= \sum_{j=0}^{k}\binom{k}{j} a^{k-j} b^j E_q\left[x^j \mathcal{I}(x \le -a/b)\right], \end{aligned} \tag{31}$$

where

$$E_q\left[x^j\mathcal{I}(x \le -a/b)\right] = \lambda\sqrt{j!}\Psi_j(-a/b,\boldsymbol{\theta}) - \sum_{n=1}^{[j/2]}\frac{(-1)^n\, j!}{n!\,(j-2n)!2^n} E_q\left[x^{j-2n}\mathcal{I}(x \le -a/b)\right], \tag{32}$$

with $\Psi_j(\cdot)$ given in (7).

Proof. It is straightforward by using (72) in Appendix 1. ■

Note that $E_q\left[x^{j-2n}\mathcal{I}(x \le -a/b)\right]$ from the right-hand side of the equation (32) is equal to $E_q\left[\mathcal{I}(x \le -a/b)\right] = Q(-a/b,\boldsymbol{\theta})$ for $j-2n=0$, and $E_q\left[x\mathcal{I}(x \le -a/b)\right] = \Psi_1(-a/b)$ for $j-2n=1$.

Corollary 4. *The equation of ϖ_4 in (27) with $z_t \sim$ i.i.d. $D(0,1)$ is given by*

$$\varpi_4 = \sum_{k=0}^{4}\psi_{4,k} E\left[\left(z_t^-\right)^k\right], \tag{33}$$

where

$$\psi_{4,0} = \beta^4 + 6\beta^2\left(\alpha_1^+\right)^2 + 4\beta\left(\alpha_1^+\right)^3 s_z + \left(\alpha_1^+\right)^4 k_z, \tag{34}$$

such that s_z *and* k_z *are given in* (10) *and* (11)*, and*

$$\begin{aligned} \psi_{4,1} &= -4\beta^3\left(\alpha_1^+ + \alpha_1^-\right), \qquad \psi_{4,2} = 6\beta^2\left[\left(\alpha_1^-\right)^2 - \left(\alpha_1^+\right)^2\right], \\ \psi_{4,3} &= 4\beta\left[\left(\alpha_1^-\right)^3 + \left(\alpha_1^+\right)^3\right], \quad \psi_{4,4} = \left(\alpha_1^-\right)^4 - \left(\alpha_1^+\right)^4. \end{aligned} \tag{35}$$

Proof. It is obtained recursively by using (27). ■

Next, we are interested in obtaining $E\left[\left(z_t^-\right)^k\right]$ under the standardized TGC distribution so as to derive the closed-form expressions of ϖ_k in (28), (30) and (33).

Corollary 5. *Consider* $z_t \sim$ *i.i.d.* $TGC\left(0,1,\boldsymbol{\theta}\right)$*, i.e.* $z_t = a + bx_t \sim$ *i.i.d.* $D\left(0,1\right)$ *such that* $x_t \sim TGC(\boldsymbol{\theta})$ *with pdf in* (5)*, then* $E\left[\left(z_t^-\right)^k\right]$ *for* $k \leq 4$ *are given by*

$$E\left(z_t^-\right) = aQ\left(-a/b,\boldsymbol{\theta}\right) + b\lambda\Psi_1\left(-a/b,\boldsymbol{\theta}\right), \tag{36}$$

$$E\left[\left(z_t^-\right)^2\right] = \left(a^2+2b^2\right)Q\left(-a/b,\boldsymbol{\theta}\right) + 2ab\lambda\Psi_1\left(-a/b,\boldsymbol{\theta}\right) + \sqrt{2}b^2\lambda\Psi_2\left(-a/b,\boldsymbol{\theta}\right), \tag{37}$$

$$\begin{aligned} E\left[\left(z_t^-\right)^3\right] &= \left(a^3+3ab^2\right)Q\left(-a/b,\boldsymbol{\theta}\right) + 3\left(a^2b+b^3\right)\lambda\Psi_1\left(-a/b,\boldsymbol{\theta}\right) \\ &\quad +3\sqrt{2}ab^2\lambda\Psi_2\left(-a/b,\boldsymbol{\theta}\right) + \sqrt{3!}b^3\lambda\Psi_3\left(-a/b,\boldsymbol{\theta}\right), \end{aligned} \tag{38}$$

and

$$\begin{aligned} E\left[\left(z_t^-\right)^4\right] &= \left(a^4+6a^2b^2+3b^4\right)Q\left(-a/b,\boldsymbol{\theta}\right) + 4\left(a^3b+3ab^3\right)\lambda\Psi_1\left(-a/b,\boldsymbol{\theta}\right) \\ &\quad +6\sqrt{2}\left(a^2b^2+b^4\right)\lambda\Psi_2\left(-a/b,\boldsymbol{\theta}\right) + 4\sqrt{3!}ab^3\lambda\Psi_3\left(-a/b,\boldsymbol{\theta}\right) \\ &\quad +\sqrt{4!}b^4\lambda\Psi_4\left(-a/b,\boldsymbol{\theta}\right). \end{aligned} \tag{39}$$

Proof. It is obtained by using (31) and (32). ■

The expressions for ϖ_k in (27) under the Normal distribution, initially obtained by FZ (2010), can be derived from the TGC density with $\theta_1 = \theta_2 = 0$ in the equations for $E\left[\left(z_t^-\right)^k\right]$ in Corollary 5, and they are shown for $k = 1, 2, 4$ in (80), see Appendix 2.

Consider $\varpi_k < 1$, then FZ (2010) provides the following result for the TGARCH model in (25):

$$E\left(\sigma_t^k\right) = \sum_{j=0}^{k}\binom{k}{j}\alpha_0^j\varpi_{k-j}E\left(\sigma_t^{k-j}\right). \tag{40}$$

Corollary 6. *The TGARCH unconditional kurtosis of* ε_t *in* (23)*, denoted as* $k_\varepsilon = E\left(\varepsilon_t^4\right)\left[E\left(\varepsilon_t^2\right)\right]^{-2}$*, such that* $\varpi_4 < 1$ *in* (33) *and* $z_t \sim i.i.d.\ D(0,1)$ *is given by*

$$k_\varepsilon = k_z \frac{E\left(\sigma_t^4\right)}{\left[E\left(\sigma_t^2\right)\right]^2} = \frac{k_z}{\left(1-\varpi_4\right)\sigma_\varepsilon^4} \sum_{j=1}^{4} \binom{4}{j} \alpha_0^j \varpi_{4-j} E\left(\sigma_t^{4-j}\right), \tag{41}$$

where $k_z = E\left(z_t^4\right)$*,* σ_ε^2 *is the equation* (29) *and* $E\left(\sigma_t^{4-j}\right)$ *is obtained recursively in* (40).

Proof. It is straightforward. ■

4.2 Power-law tail property

Suppose there exists a positive real number $\zeta > 0$ such that $E\left(c_t^\zeta\right) = 1$ with c_t in (26). According to the theory in Kesten (1973), the stationary solution of the stochastic difference equation (25) follows a heavy-tailed distribution:

$$P\left\{\sigma_t > x\right\} \sim Ax^{-\zeta} \quad \text{as} \quad x \to \infty,$$

where ζ is the tail index of σ_t and $A > 0$ is the tail scale.[4] Suppose that $E\left(|z_t|^{\zeta+\epsilon}\right) < \infty$ for some $\epsilon > 0$, then Mikosch and Starica (2000) derive the following result:

$$P\left\{|\varepsilon_t| > x\right\} = P\left\{|\sigma_t z_t| > x\right\} \sim E\left(|z_t|^\zeta\right) P\left\{\sigma_t > x\right\} \quad \text{as} \quad x \to \infty.$$

In short, $|\varepsilon_t|$ has a similar tail behavior as σ_t, i.e. the tail index of $|\varepsilon_t|$ equals ζ. For the existence of the p-th moment of ε_t, it must be verified that $E\left(|\varepsilon_t|^p\right) < \infty$. Since the value of ζ characterizes the maximal order of finite moments of ε_t, then

$$E\left(|\varepsilon_t|^p\right) < \infty \quad \text{if} \quad p < \zeta \quad \text{and} \quad E\left(|\varepsilon_t|^p\right) = \infty \quad \text{if} \quad p \geq \zeta. \tag{42}$$

According to (42), it is verified that $E\left(|\varepsilon_t|\right) < \infty$ if only if $\zeta > 1$. The second moment $E\left(\varepsilon_t^2\right) < \infty$, and thus $\sigma_\varepsilon^2 < \infty$, if and only if $\zeta > 2$. In short, consider $z_t \sim i.i.d.\ D(0,1)$, then $\varpi_1 < 1$ in (28) if $\zeta > 1$ and $\varpi_2 < 1$ in (30) if $\zeta > 2$. Similarly, the fourth moment $E\left(\varepsilon_t^4\right) < \infty$, and hence kurtosis $k_\varepsilon < \infty$ in (41), if and only if $\zeta > 4$. This condition implies $\varpi_4 < 1$ in (33). Finally, the particular case of ϖ_k in (27) for $k = 1, 2, 4$ with $z_t \sim i.i.d.\ TGC(0, 1, \boldsymbol{\theta})$ is obtained using Corollary 5.

4.3 Conditional log-likelihood

The conditional density of r_t can be expressed in terms of the conditional pdf of x_t in (4) as

$$f\left(r_t \,|I_{t-1}\right) = \frac{1}{b_t \sigma_t} q\left(\frac{z_t - a_t}{b_t} \,|I_{t-1}\right), \tag{43}$$

[4] See Gabaix (2009) for an introduction to power-law distributions.

where $a_t = a(\boldsymbol{\theta}_t)$, $b_t = b(\boldsymbol{\theta}_t)$, $\boldsymbol{\theta}_t = (\theta_{1t}, \theta_{2t})$ is measurable with respect to the information set I_{t-1} and $\theta_{it} = \theta_{it}(\boldsymbol{\vartheta}_i)$ such that $\boldsymbol{\vartheta}_i$ is the parameter vector underlying the dynamics of θ_{it} as can be seen in section 6. The log-likelihood function corresponding to a particular observation r_t, denoted as l_t, is given as

$$\begin{aligned} l_t &= -\ln \sigma_t(\boldsymbol{\varphi}) - \ln b(\boldsymbol{\theta}_t) - \frac{1}{2}\ln(2\pi) + \ln \lambda(\boldsymbol{\theta}_t) \\ &\quad -\frac{1}{2}\left(\frac{z_t(\boldsymbol{\varphi}) - a(\boldsymbol{\theta}_t)}{b(\boldsymbol{\theta}_t)}\right)^2 + \ln\left[\psi\left(\frac{z_t(\boldsymbol{\varphi}) - a(\boldsymbol{\theta}_t)}{b(\boldsymbol{\theta}_t)}\right)\right]^2, \end{aligned} \tag{44}$$

where $z_t(\boldsymbol{\varphi}) = (r_t - \mu_t(\boldsymbol{\varphi}))/\sigma_t(\boldsymbol{\varphi})$ and $\boldsymbol{\varphi}$ is the parameter vector to model both the conditional mean and variance. We will consider in our empirical application in section 5 the following case: (i) $\mu_t = \mu$, σ_t in (24), and hence $\boldsymbol{\varphi} = \left(\mu, \alpha_0, \beta, \alpha_1^+, \alpha_1^-\right)$, and (ii) $\boldsymbol{\theta}_t = \boldsymbol{\theta}$. Finally, this particular case means that (43) can be replaced by the expression: $f(r_t|I_{t-1}) = \frac{1}{b\sigma_t} q\left(\frac{z_t - a}{b}\right)$.

4.4 Conditional asymmetric risk measures

Let $F(r_t|I_{t-1})$ denote the cdf of r_t with the corresponding pdf in (43),

$$F(r_t|I_{t-1}) = \int_{-\infty}^{r_t} f(u|I_{t-1})\,du = Q\left(\frac{r_t - \kappa_{0t}}{\kappa_{1t}}|I_{t-1}\right), \tag{45}$$

where $Q(\cdot|I_{t-1})$ is the conditional cdf of $Q(\cdot)$ in (8) and both $\kappa_{0t} = \mu_t + a_t\sigma_t$ and $\kappa_{1t} = b_t\sigma_t$ are measurable respecting I_{t-1}. The conditional α-quantile (or VaR) of the stock return r_t is given by $r_{\alpha,t} = F^{-1}(\alpha|I_{t-1})$. Then,

$$r_{\alpha,t} = \kappa_{0t} + \kappa_{1t} Q_t^{-1}(\alpha), \tag{46}$$

such that $Q_t^{-1}(\alpha) = \inf\{x\,|Q(x|I_{t-1}) \geq \alpha\}$ is the conditional α-quantile of x_t with $q(\cdot|I_{t-1})$ as pdf.

The conditional ES of r_t is easily computed as

$$\begin{aligned} ES_t(\alpha) &= E_{t-1}(r_t|r_t \leq r_{\alpha,t}) \\ &= \kappa_{0t} + \kappa_{1t} E_{t-1}(x_t|x_t \leq x_{\alpha,t}), \end{aligned} \tag{47}$$

where $E_{t-1}(x_t|x_t \leq x_{\alpha,t})$ is the conditional version of (13) and $x_{\alpha,t} = (r_{\alpha,t} - \kappa_{0t})/\kappa_{1t}$ with $r_{\alpha,t}$ as the VaR in (46). Note that $E_{t-1}(\cdot)$ denotes the shortening of $E(\cdot|I_{t-1})$.

The conditional LPM of order 1 and τ as the threshold of r_t is given by

$$\begin{aligned} LPM_t(\tau, 1) &= \int_{-\infty}^{\tau} (\tau - r_t) f(r_t|I_{t-1})\,dr_t \\ &= (\tau - \kappa_{0t}) Q_t(\tau_t) - \kappa_{1t} Q_t(\tau_t) E_{t-1}(x_t|x_t \leq \tau_t), \end{aligned} \tag{48}$$

where $Q_t(\cdot)$ denotes $Q(\cdot\,|I_{t-1})$ and $\tau_t = (\tau - \kappa_{0t})/\kappa_{1t}$. Finally, the conditional LPM of order 2 is

$$\begin{aligned} LPM_t(\tau,2) &= \int_{-\infty}^{\tau} (\tau - r_t)^2 f(r_t\,|I_{t-1})\, dr_t \\ &= (\tau-\kappa_{0t})^2 Q_t(\tau_t) - 2(\tau-\kappa_{0t})\kappa_{1t} Q_t(\tau_t) E_{t-1}(x_t\,|x_t \le \tau_t) \\ &\quad + \kappa_{1t}^2 E_{t-1}\left[x_t^2 \mathcal{I}(x_t \le \tau_t)\right]. \end{aligned} \tag{49}$$

5 Empirical application

5.1 Dataset and summary statistics

The data used are daily percent log returns computed as $r_t = 100 \ln(P_t/P_{t-1})$ from samples of daily closing prices $\{P_t\}_{t=1}^{T}$ for Eurostoxx50 and Nikkei indexes, Japanese Yen to U.S. dollar (JAP-US) and U.S. dollar to pound sterling exchange rates (US-UK) and West Texas Intermediate Crude Oil, all obtained from the New York Stock Exchange, sampled from January 14, 1999 to January 14, 2019 for a total of $T = 5,218$ observations. We also consider Bitcoin prices sampled from from July 18, 2010 to July 31, 2018 ($T = 2,936$). All data series were obtained from Datastream, apart from Bitcoin series downloaded from coindesk.com. Table 1 exhibits summary statistics of all data returns. Clearly, all the series show high leptokurtosis with the Bitcoin presenting the largest kurtosis (14.96) and the Oil the smallest (7.23). The skewness is negative in all series, with the largest (in absolute value) corresponding to US-UK (-0.58) and the smallest to the Eurostoxx (-0.08). In all cases, the Jarque-Bera (J-B) test rejects the null of normality, motivating the use of alternative distributions to the Gaussian for modeling returns.

Table 1: Summary statistics for daily percent log returns

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
Mean	0.008	-0.002	-0.001	-0.005	0.027	0.389
Median	0.017	0.016	0.005	0.000	0.000	0.203
Max	11.644	11.965	3.710	4.474	16.414	42.458
Min	-11.186	-11.102	-6.582	-8.312	-17.092	-49.144
Std. dev.	1.457	1.560	0.684	0.582	2.375	5.735
Skewness	-0.235	-0.085	-0.473	-0.583	-0.170	-0.320
Kurtosis	7.445	8.896	8.096	14.482	7.235	14.967
J-B stat	4363	7564	5839	28953	3924	17565
Observations	5217	5217	5217	5217	5217	2935

This table presents the summary statistics for daily percent log returns. The Jarque-Bera (J-B) statistic is asymptotically distributed as a Chi-square with two degrees of freedom, χ^2_2. The critical value of χ^2_2 for the 5% significance level is 5.99. The sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from January 15, 1999 to January 14, 2019; and for Bitcoin returns is from July 19, 2010 to July 31, 2018.

5.2 Estimation results

Initially, we adopt some density functions with constant parameters across time for the conditional standardized returns in (23), i.e. $z_t \sim i.i.d.\ D(0,1)$. We consider the following cases: (i) the standard Normal distribution (N hereafter), i.e. $D(0,1) = N(0,1)$; (ii) the GC symmetric density, proposed by Zoia et al. (2018), which gathers positive excess kurtosis (GCK hereafter), i.e. $D(0,1) = GC(0, ek)$; and (iii) $D(0,1)$ is the standardized TGC distribution denoted as $TGC(0,1,\boldsymbol{\theta})$, i.e. $z_t = a + bx_t$ such that $x_t \sim TGC(\boldsymbol{\theta})$ with pdf in (4). The conditional mean and volatility of r_t in (23) are given by $\mu_t = \mu$ and σ_t in (24).

A comparative analysis for the models in-sample goodness-of-fit is also carried out through likelihood ratio (LR) tests. In particular, we employ Vuong's LR test (LRV) (1989) for two nonnested i and j models. The LRV statistic is based on the null hypothesis of being the two candidate models equally close to the true specification, and it is defined as

$$LRV_{i,j} = T^{-1/2}\left(LL_i - LL_j\right)/s_T \xrightarrow{d} N(0,1), \tag{50}$$

such that LL_k denotes the log-likelihood value for model k, and s_T is the sample standard deviation obtained in the usual manner, i.e.

$$s_T^2 = T^{-1}\sum_{t=1}^{T}\left(\ln\left(l_t^i/l_t^j\right)\right)^2 - \left(T^{-1}\sum_{t=1}^{T}\ln(l_t^i/l_t^j)\right)^2, \tag{51}$$

where l_t^k denotes the log-likelihood corresponding to a particular observation t for model k.

Table 2 presents the estimation of the TGARCH model under TGC, GCK and Normal (henceforth, TGC-TGARCH, GCK-TGARCH and N-TGARCH) for the whole sample of each return series. A first observation is that the TGARCH parameter estimates are very similar under any of the three considered densities and so, to simplify the presentation, only the estimates under the TGC-TGARCH are reported in the table. These estimates show that the returns for all assets exhibit volatility clustering and asymmetric response to good and bad news. Specifically, we find that for the Bitcoin series the TGARCH parameter estimates differ significantly from those of the other series reflecting rather different volatility patterns. Second, both TGC and GCK density parameters are significant for all series.[5] Note that the GCK parameters are not reported for the sake of simplicity. Third, all series except Bitcoin verify the second-order stationarity condition, i.e. the estimate of ϖ_2 in (30) is lower than one, for both TGC and GCK. It is also verified that the unconditional standard deviations, denoted as σ_ε in (29), are close to the sample ones in Table 1 for the two distributions. For instance, the estimation of σ_ε under TGC-TGARCH is equal to 1.723 (1.507 under GCK-TGARCH) and its sample standard deviation is 1.457 for Nikkei. Respecting the N-TGARCH (not exhibited in Table 2), it is verified that $\varpi_2 < 1$ for all series including Bitcoin but the estimation of σ_ε is around 21, which is far away from the sample standard deviation of 5.7, suggesting a poor fit for Bitcoin. The N-TGARCH estimations of σ_ε for the other return series are worse than those under TGC and GCK when comparing to the sample standard deviations in Table 1.

[5] We have also estimated the GC with pdf in (1) and found similar parameter estimates as the ones in Table 2 for the TGC.

In regard to the TGARCH models' fit, the TGC yields higher log-likelihood (LL) values than the GCK and Normal for all return series. The LR tests for TGC and GCK versus the Normal, not reported but available, reject the Normal as the null at any significance level. The results of the LRV test in (50) for TGC and GCK (denoted as i and j, respectively) show that the TGC provides statistically significant better in-sample goodness-of-fit than the GCK for all return series, except for the Bitcoin for which, although the LL values are higher under the TGC, the difference is not statistically significant. The Akaike information criteria (AIC) reinforce the LR tests. Lower AIC values indicate better goodness-of-fit, AIC_N is higher than both AIC_{GCK} and AIC_{TGC}.

Finally, we provide a robustness check for the specification of the conditional variance. Thus, we analyze whether there are differences in the fit of the TGC pdf when the conditional variance is modelled through the NAGARCH as an alternative to the TGARCH. Our results, see Table A3 in Appendix 3, show that the TGC-NAGARCH estimation gathers very similar conditional variance dynamics than the TGC-TGARCH in Table 2. If anything, we observe a slightly lower persistence (lower value of β) for Nikkei and Eurostoxx50. The NAGARCH second-order stationarity condition is also not verified for Bitcoin, i.e. $\varpi_1 > 1$ in (83), see Appendix 2. The LRV test results show that differences for the NAGARCH under i =TGC and j =GCK in (50) are not statistically significant.

5.3 Tail-index estimation

We employ the methodology of Gabaix and Ibragimov (2011) to obtain robust estimates of the tail index ζ together with its confidence interval at 95%. Remember that $\varpi_2 < 1$ ($\varpi_1 < 1$) in TGARCH (NAGARCH) model is equivalent to $\zeta > 2$. Our results, presented in Table A4 in Appendix 4, show that the null $H_0 : \zeta = 2$ is rejected for all return series in favor of the one-sided alternative $H_a : \zeta > 2$ for both 5% and 10% truncation levels. In short, we find evidence in favor of finite unconditional variance for all return series, although there is a contradiction for Bitcoin since $\varpi_2 > 1$ under TGARCH in Table 2 and $\varpi_1 > 1$ under NAGARCH in Table A3 in Appendix 3. We also study the finiteness of the unconditional third and fourth moments of such returns. For all series except Bitcoin the null $H_0 : \zeta = 4$ is not rejected at 5% truncation level, although we cannot reject $H_0 : \zeta = 4$ in favour of $H_a : \zeta > 4$. Respecting the 10% truncation level, $H_0 : \zeta = 4$ is always rejected. Note that the estimation of ϖ_2 under the NAGARCH does only report values of ϖ_2 lower than one for JAP-US and US-UK. Respecting testing $H_0 : \zeta = 3$, we can reject it in favour of $H_a : \zeta > 3$ for Nikkei at both truncation levels, JAP-US at 5% truncation level and US-UK at 10%.

Finally, the implied tail-index estimation for asymmetric GARCH models could be obtained numerically through simulations from Kesten's equation following the studies in Chan, Li, Peng and Zhang (2013) for ARCH; and Zhang, Li and Peng (2019) for GARCH.[6] However, this issue is beyond the scope of this paper. Nonetheless, in Appendix 5 we provide an advance on this research line for the case of NAGARCH. Table A5 presents results of a sensitivity analysis of the tail-index estimation through Kesten's equation for different values of ϖ_2, see equation (84) in Appendix 2, according to alternative values of the excess kurtosis parameter under the GC pdf in (1). As expected, we find that the higher ϖ_2 the lower the tail index. In short, our

[6] We are indebted to the referee for suggesting this point.

results do agree with the empirical evidence of the non-existence of higher-order unconditional moments (skewness or kurtosis, or both) for financial returns series, see Ibragimov, Ibragimov and Walden (2015), and references therein.

5.4 Backtesting the density

For all the backtesting procedures, 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. Indeed, we adopt a two-stage estimation method to each window as can be seen, among others, in Zhu and Galbraith (2011). The mean and TGARCH parameters are estimated by quasi-maximum likelihood (QML), then the TGC and GCK density parameters are obtained by ML using the standardized residuals, z_t, from the first stage.

5.4.1 P-value discrepancy plots

First, we test the density forecasting performance of the models following the methodology in Diebold, Gunther and Tay (1998). The application of this methodology is based on the cdf evaluated at the one-step-ahead realized returns through the OOS period. The resulting so-called probability integral transforms (PIT) sequences verify that $\{u_t\}_{t=1}^{N} \sim i.i.d.\ U(0,1)$ under the correct one-step ahead cdf specification with $u_t = D\left(r_t \left| I_{t-1} \right.\right)$ where $D\left(\cdot \left| I_{t-1} \right.\right)$ denotes a conditional cdf. We use the p-value plot methods in Davidson and MacKinnon (1998) applied to compare models forecasting performance. Thus, if the model is correctly specified the difference between the cdf of u_t and the 45 degree line should tend to zero. The empirical distribution function of u_t can be easily computed as

$$\widehat{P}_{p_t}(y_\varrho) = \frac{1}{N}\sum_{t=1}^{N} \mathcal{I}(u_t \leq y_\varrho), \tag{52}$$

where $\mathcal{I}(u_t \leq y_\varrho)$ is an indicator function and y_ϱ is an arbitrary grid of ϱ points.[7] Alternatively, the p-value discrepancy plot (i.e. plotting $\widehat{P}_{p_t}(y_\varrho) - y_\varrho$ against y_ϱ) can be more revealing when it is necessary to discriminate among specifications that perform similarly in terms of the p-value plot. Consequently, under correct density specification, the variable $\widehat{P}_{p_t}(y_\varrho) - y_\varrho$ must be close to zero.

Figure 6 presents in panel 1 the p-value discrepancy plots for all models and series. A first observation that emerges from the plots is that the TGARCH models under TGC and GCK perform overall better than the Normal. We find that PITs from the TGC depart further from those of the GCK and Normal for mid quantiles (center of the distribution). It is worth noting the clustering of the GCK's plots with those of the Normal in the center of the distribution. With the exception of the Bitcoin series, for which there are no visually noticeable differences between the TGC and GCK, the former provides better fit of the left-tail of the distribution for the rest of the series, as can be seen in panel 2 (zoom in panel 1 for the left-tail) by the smaller distance of the TGC plots to the x-axis.

[7] We use the following $\varrho = 215$ points grid: $y_\varrho \in \{0.001, 0.002, ..., 0.01, 0.015, ..., 0.99, 0.991, ..., 0.999\}$, since it highlights the goodness-of-fit in the distribution tails.

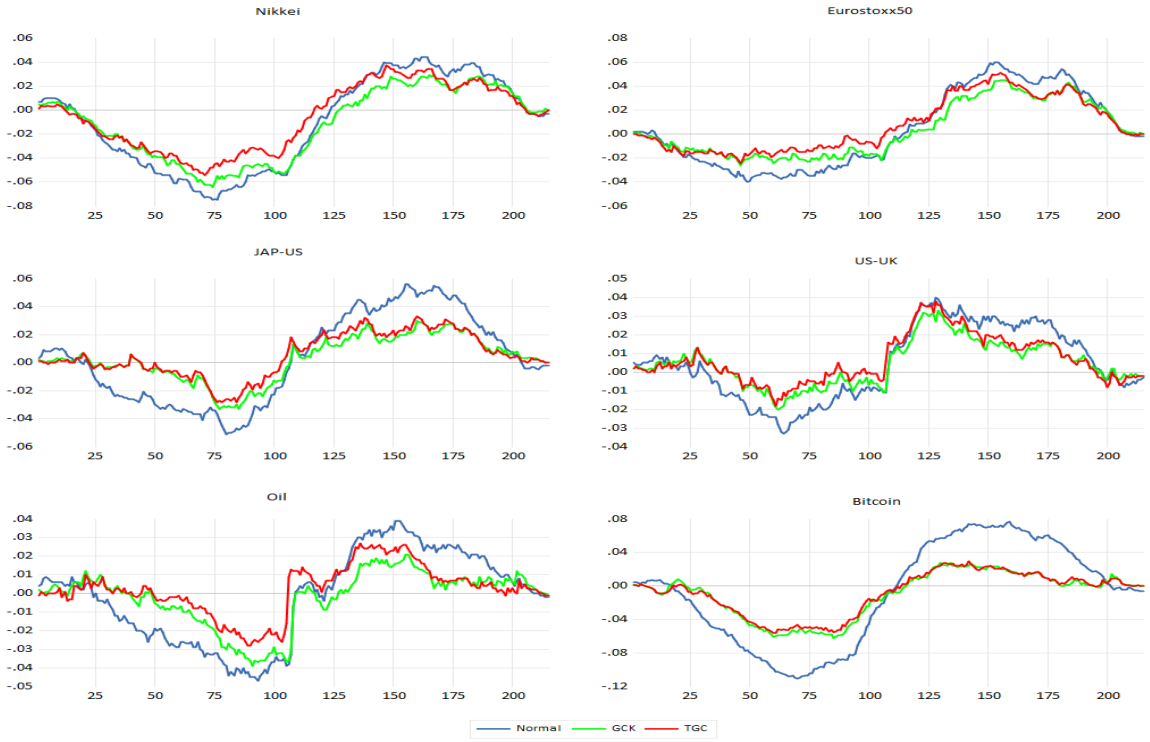
Table 2: Estimation results of TGC-TGARCH model

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
μ	-0.0091	0.0016	0.0031	-0.0018	0.0168	0.2407*
	(0.017)	(0.015)	(0.009)	(0.007)	(0.028)	(0.102)
α_0	0.0353*	0.0271*	0.0078*	0.0043*	0.0163*	0.3566*
	(0.009)	(0.005)	(0.003)	(0.002)	(0.005)	(0.164)
β	0.9155*	0.9218*	0.9483*	0.9494*	0.9572*	0.7168*
	(0.012)	(0.009)	(0.008)	(0.001)	(0.007)	(0.063)
α_1^+	0.0333*	0.0162***	0.0490*	0.0504*	0.0234*	0.3484*
	(0.009)	(0.009)	(0.008)	(0.018)	(0.007)	(0.074)
α_1^-	0.1332*	0.1359*	0.0597*	0.0631*	0.0729*	0.3717*
	(0.017)	(0.014)	(0.011)	(0.012)	(0.012)	(0.082)
θ_1	-0.0653*	-0.0538*	-0.0386*	-0.0289***	-0.0806*	-0.0231
	(0.018)	(0.017)	(0.018)	(0.018)	(0.017)	(0.033)
θ_2	0.3134*	0.2880*	0.4649*	0.2864*	0.3813*	0.9846*
	(0.034)	(0.035)	(0.035)	(0.038)	(0.035)	(0.053)
ϖ_2	0.9657	0.9768	0.9809	0.9887	0.9921	1.0147
σ_ε	1.7230	1.8002	0.7745	0.6934	3.3806	- - - -
ϖ_4	0.9652	0.9924	0.9669	0.9828	0.9964	- - - -
LL_{TGC}	-8738.13	-8654.96	-4952.65	-4049.52	-11235.17	-4615.58
LL_{GCK}	-8751.06	-8666.09	-4967.91	-4054.59	-11260.13	-4618.38
LL_N	-8806.89	-8713.08	-5094.35	-4097.99	-11348.96	-4788.81
LRV	2.6106*	2.5366*	3.5232*	1.6842**	4.5188*	1.0543
AIC_{TGC}	3.3512	3.3193	1.9000	1.5538	4.3085	3.1476
AIC_{GCK}	3.3560	3.3234	1.9057	1.5555	4.3179	3.1492
AIC_N	3.3772	3.3412	1.9539	1.5720	4.3517	3.2649

Model specification: $r_t = \mu + \varepsilon_t,\ \ \varepsilon_t = \sigma_t z_t,\ \ \ z_t \sim i.i.d.\ TGC\left(0, 1, \boldsymbol{\theta}\right),\ \ \sigma_t = \alpha_0 + \beta\sigma_{t-1} + \alpha_1^+\varepsilon_{t-1}^+ - \alpha_1^-\varepsilon_{t-1}^-$.

This table presents ML estimates of parameters of the TGC-TGARCH model for percent log return series. The sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from January 15, 1999 to January 14, 2019 (total: 5,217 obs.); and for Bitcoin returns it is from July 19, 2010 to July 31, 2018 (total: 2,935 obs.). Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*) and (**) indicate significance at 1% and 5% levels, respectively. The second-order and fourth-order moment stationarity conditions of TGC-TGARCH must verify $\varpi_k < 1$, $k = 2, 4$, see equations (30), (33) and Corollary 5. Let σ_ε denote the TGC-TGARCH unconditional standard deviation of ε_t. AIC denotes the Akaike information criterion (the lower the AIC value, the better the goodness-of-fit). LL_j denotes the log-likelihood value for j-TGARCH model where $j = TGC, GCK, N$. Finally, LRV denotes the nonnested LR test statistic of Vuong (1989) in (50) for $i = TGC$ and $j = GCK$ under TGARCH model.

Panel 1: Full density



Panel 2: Left 10% tail

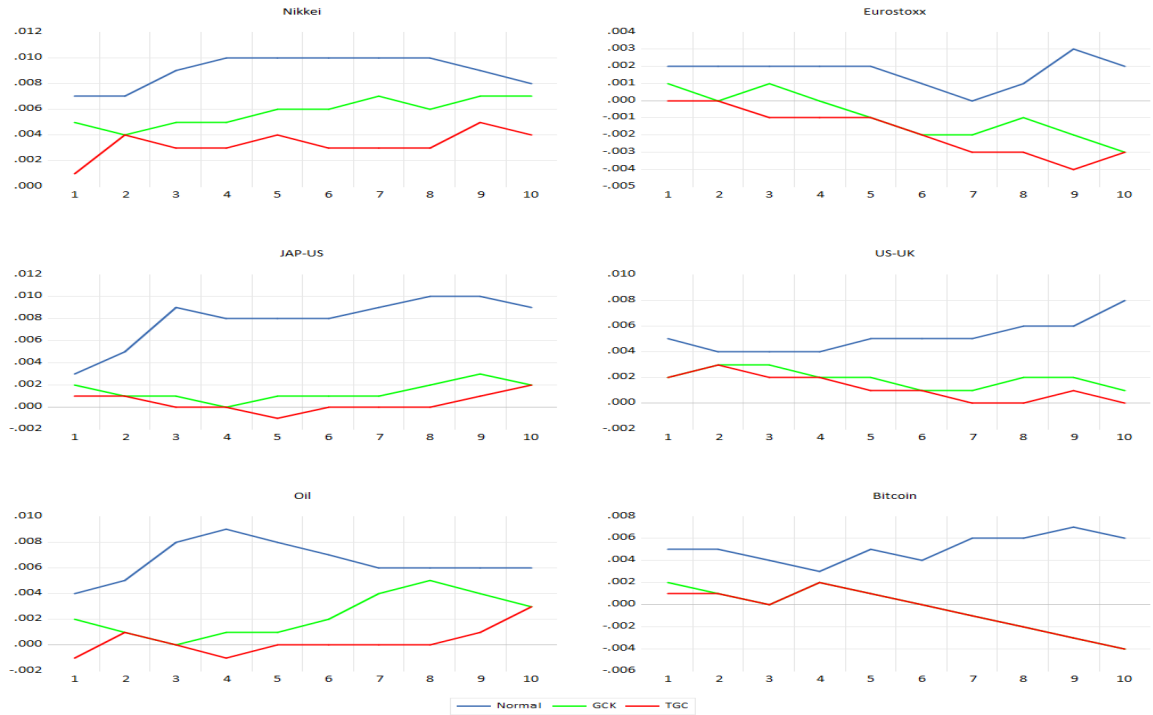


Figure 6: P-value discrepancy plots for PIT series obtained from TGARCH model under TGC, GCK and Normal pdfs. Return series: Nikkei, Eurostoxx50, JAP-US, US-UK, Oil and Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns is from November 4, 2015 to July 31, 2018. Predictions: 1000.

5.4.2 Proper scoring rules

Second, to evaluate the relative density forecasting performance of the models we use strictly proper scoring rules described in Amisano and Giacomini (2007). A scoring rule is a loss function $\Upsilon(\widetilde{f}_t, r_t)$ whose arguments are the density forecast $\widetilde{f}_t = f_{t-1}(r_t)$ and the future realization of the return, r_t. In this paper, we use the weighted logarithmic scoring rule:

$$\Upsilon_j(\widetilde{f}_t, z_t) = \omega_j(z_t) \ln \widetilde{f}_t. \tag{53}$$

This is a strictly proper scoring rule that rewards accurate density forecasts by setting a high probability to the event that actually occurred. The weight functions $\omega_1(z_t) = \phi(z_t)$, $\omega_2(z_t) = \Phi(z_t)$ and $\omega_3(z_t) = 1 - \Phi(z_t)$ emphasize, respectively, the center, the right-tail and the left-tail. The density forecast models can be ranked by comparing their average scores:

$$\overline{\Upsilon}_j(\widetilde{f}_t, z_t) = N^{-1} \sum_{t=1}^{N} \Upsilon_{jt}(\widetilde{f}_t, z_t). \tag{54}$$

So, we prefer model f if $\overline{\Upsilon}_j(\widetilde{f}_t, z_t) > \overline{\Upsilon}_j(\widetilde{g}_t, z_t)$, and prefer model g otherwise. The null hypothesis $H_0 : E\left[\overline{\Upsilon}_j(\widetilde{f}_t, z_t) - \overline{\Upsilon}_j(\widetilde{g}_t, z_t)\right] = 0$ is tested through the Diebold and Mariano (1995) (DM hereafter) test.

Table 3: Density Forecasting

		Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
TGC	left	-0.5768	-0.6871	-0.6957	-0.7686	-0.6957	-0.5768
	center	-0.2562	-0.3265	-0.3082	-0.3191	-0.3082	-0.2562
	right	-0.5773	-0.6354	-0.6424	-0.6915	-0.6424	-0.5773
GCK	left	-0.6394	-0.6919	-0.7124	-0.7762	-0.7124	-0.6393
	center	-0.2953	-0.3312	-0.3185	-0.3223	-0.3185	-0.2953
	right	-0.6456	-0.6484	-0.6616	-0.6982	-0.6616	-0.6456
N	left	-0.6785	-0.7011	-0.7308	-0.7890	-0.7308	-0.6785
	center	-0.3371	-0.3396	-0.3337	-0.3295	-0.3337	-0.3371
	right	-0.6789	-0.6589	-0.6753	-0.7069	-0.6753	-0.6789

This table presents the results of average logarithmic scores in (54) for one-step-ahead density forecast from TGARCH model under TGC, GCK and Normal (N) pdfs. Return series: Nikkei, Eurostoxx50, JAP-US, US-UK, Oil and Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns is from November 4, 2015 to July 31, 2018. Predictions: 1000.

Table 3 presents the results of the weighted average scores for the alternative TGARCH models. A first observation that stands out is that all TGC's weighted scores are higher than those of the GCK and Normal for all series. The Normal provides systematically the lowest scores. Besides, differences in models' scores are statistically significant as the DM test rejects the null of equal score for all cases. These results show that the TGC provides the statistically significant better performance to forecast the tails and the center of the return densities.

5.5 Backtesting VaR and ES

We evaluate the forecasting performance for the left-tail of the return conditional distribution under alternative density specifications for z_t. Given a nominal coverage rate α, the one-day TGC-TGARCH VaR is

$$VaR_t(\alpha) = \kappa_{0,t} + \kappa_{1,t} Q^{-1}(\alpha), \tag{55}$$

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. The quadratic loss function, proposed by López (1999), incorporates the exception magnitude and so, it provides useful information to discriminate among similar models according to the unconditional coverage test, and it is given by

$$QL_t(\alpha) = (r_t - VaR_t(\alpha))^2 \times h_t(\alpha). \tag{56}$$

The sample average of (56) for the OOS period of N days is

$$AQL(\alpha) = N^{-1} \sum_{t=1}^{N} QL_t(\alpha). \tag{57}$$

5.5.1 Backtesting VaR

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 following test statistics, proposed by Kupiec (1995), which converges asymptotically to a standard Normal distribution, i.e.

$$U_{VaR}(\alpha) = \frac{\sqrt{N}\left(\overline{h}(\alpha) - \alpha\right)}{\sqrt{\alpha(1-\alpha)}} \overset{a}{\sim} N(0,1), \tag{58}$$

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 $u_t = F(r_t | I_{t-1})$ in (45). For testing the null hypothesis for the conditional backtest, $H_{0,C} : E[h_t(\alpha) - \alpha | I_{t-1}] = 0$, we implement the approach by Escanciano and Olmo (2010) based on the Box-Pierce test statistic:

$$C_{VaR}(m) = N \sum_{i=1}^{m} \widehat{\rho}_j^2 \overset{a}{\sim} \chi_m^2, \tag{59}$$

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

$$\hat{\gamma}_j = \frac{1}{N-j} \sum_{t=1+j}^{N} \left(\widehat{h}_t(\alpha) - \alpha\right)\left(\widehat{h}_{t-j}(\alpha) - \alpha\right). \tag{60}$$

5.5.2 Backtesting ES

The unconditional and conditional ES backtests are the analogues of 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

$$\begin{aligned}\mathcal{H}_t(\alpha) &= \int_0^{\alpha} h_t(u)\,du \\ &= \frac{1}{\alpha}(\alpha - u_t)\,\mathcal{I}(u_t \leq \alpha). \end{aligned} \tag{61}$$

Note that the equation (61) measures the distance of the returns from the corresponding α-quantile in (55) during the violations. It is shown that $\{\mathcal{H}_t(\alpha) - \alpha/2\}_{t=1}^{\infty}$ follows the MDS property. Thus, 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

$$U_{ES} = \frac{\sqrt{N}\left(\overline{\mathcal{H}}(\alpha) - \frac{\alpha}{2}\right)}{\sqrt{\alpha\left(\frac{1}{3} - \frac{\alpha}{4}\right)}} \overset{a}{\sim} N(0,1), \tag{62}$$

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 corresponding test statistics the Box-Pierce one given by

$$C_{ES}(m) = N\sum_{i=1}^{m}\hat{\rho}_j^2 \overset{a}{\sim} \chi_m^2, \tag{63}$$

such that $\hat{\rho}_j = \hat{\gamma}_j/\hat{\gamma}_0$ is the j-th lag of the sample autocorrelation with

$$\hat{\gamma}_j = \frac{1}{N-j}\sum_{t=1+j}^{N}\left(\widehat{\mathcal{H}}_t(\alpha) - \frac{\alpha}{2}\right)\left(\widehat{\mathcal{H}}_{t-j}(\alpha) - \frac{\alpha}{2}\right). \tag{64}$$

5.5.3 Backtesting results

Following Kerkhof and Melenberg (2004) and others, a larger coverage level α for ES than VaR is selected to compare both risk measures. The coverage level for ES is twice (or close to twice) than that of VaR, indeed. Table 4 shows the results of the descriptive analysis of violations. First, we find that for the low coverage levels suggested by the Basel Committee (i.e. VaR(1%), ES(2.5%)) the TGC-TGARCH performs much better than the other TGARCH models for all series except for Bitcoin, where both TGC and GCK provide rather the same performance and beat the Normal. Note that this evidence is in line with what we can infer from Figure 6, if we study the left-tail risk by considering only the first ten discrepancy p-value points exhibited in panel 2, where the TGC p-value discrepancy points are the closest to the x-axis for all series. In short, the skewness implied in the TGC contribute to a better fit, respecting the symmetric GCK, of the very far left-tail of the return distributions. Second, for higher coverage levels, i.e. VaR(5%) and ES(10%), the TGC is always better than the GCK. Nevertheless, for JAP-US and Bitcoin series the Normal beats the TGC. Third, we report the significance at five percent level of both unconditional and conditional backtesting for VaR and ES. We set $m = 5$ in (59) and (63). It is verified that the null hypotheses are accepted for JAP-US, US-UK and Bitcoin for both TGC-TGARCH and GCK-TGARCH. Note that for Eurostoxx50 the null of conditional backtest for VaR(5%) is always rejected, and the same goes for Oil for ES(10%). We also find many rejections under the N-TGARCH.

Finally, the previous VaR results are reinforced by the magnitude of exceptions for VaR measured through the AQL statistic in (57) and presented in Table 5. First, we find that the TGC (Normal) yields always

the lowest (highest) AQL values for the 1% level. Note that only for Bitcoin the AQL value under GCK is slightly lower than under TGC. Second, for the 5% level, the Normal provides a better performance than i) GCK in all series, and ii) TGC for JAP-US and Bitcoin.

Table 4: Descriptive analysis of violations

	VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)	VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)
	Nikkei				Eurostoxx50			
TGC	14	15.38	44	43.21	7	8.36	40^{c}	37.89^{u}
GCK	18^{u}	17.40	46	45.75	7	9.68	41^{c}	40.17
N	18^{u}	19.40^{u}	46	45.26	12	11.49	41^{c}	38.90^{u}
	JAP-US				US-UK			
TGC	12	13.77	53	50.22	11	14.98	53	53.71
GCK	12	14.82	54	51.34	12	16.40	56	55.26
N	19^{u}	20.60^{u}	52	49.48	18^{u}	19.56^{u}	54	53.81
	Oil				Bitcoin			
TGC	13	11.97	52	52.06^{c}	6	7.66	48	45.09
GCK	13	14.21	56^{c}	55.35^{c}	6	7.48	54	47.85
N	16	18.36	54^{c}	53.45^{c}	16	18.09	47	44.48

This table shows the violations for VaR and the cumulative violations given in (61) for ES from TGARCH model under TGC, GCK and Normal (N) pdfs. We also reports the significance for (i) the VaR backtesting tests in (58) and (59), and (ii) the ES backtesting tests in (62) and (63). We set $m = 5$ in the Box-Pierce test statistic for the two conditional backtests. The superscripts u and c indicate significance at five percent level for the unconditional and conditional backtests, respectively. Return series: Nikkei, JAP-US, Oil, US-UK, Eurostoxx50, Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns is from November 4, 2015 to July 31, 2018. Predictions: 1000.

Table 5: Average quadratic loss (AQL)

	VaR(1%)	VaR(5%)	VaR(1%)	VaR(5%)	VaR(1%)	VaR(5%)
	Nikkei		Eurostoxx50		JAP-US	
TGC	0.0186	0.0725	0.0554	0.0925	0.0057	0.0213
GCK	0.0229	0.0772	0.0585	0.0951	0.0062	0.0218
N	0.0279	0.0761	0.0625	0.0943	0.0089	0.0211
	US-UK		Oil		Bitcoin	
TGC	0.0432	0.0599	0.0135	0.1277	0.1299	0.7575
GCK	0.0439	0.0605	0.0193	0.1402	0.1292	0.8524
N	0.0457	0.0601	0.0326	0.1345	0.2872	0.7235

This table exhibits the results of the AQL in (57) for VaR from TGARCH model under TGC, GCK and Normal (N) pdfs. Return series: Nikkei, JAP-US, Oil, US-UK, Eurostoxx50, Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns is from November 4, 2015 to July 31, 2018. Predictions: 1000.

5.6 Comparative analysis

This section provides a comparison of the TGC distribution and the popular skewed-t (ST) distribution of Hansen (1994) with parameters $\delta \in (-1, 1)$ and $\upsilon > 2$ that control for skewness and kurtosis, respectively.[8] Table 6 reports the estimates of the ST-TGARCH model. To avoid duplicity, we do not display the TGARCH estimates as they are very similar to those of the TGC-TGARCH in Table 2. The estimation results show that both υ and δ are significant, except the skewness parameter δ for Bitcoin. The results from the LRV test in (50) for both ST and TGC pdfs under the TGARCH, i.e. $T^{-1/2}\left(LL_{ST} - LL_{TGC}\right)/s_T \xrightarrow{d} N\left(0,1\right)$, show that for Nikkei, Eurostoxx50 and Bitcoin, the TGC provides better fit than the ST, although differences are not statistically significant for Bitcoin. Whilst for JAP-US, US-UK and Oil, the ST provides better fit, with differences being statistically significant only for JAP-US.

The p-value discrepancy plots in Figure 7 show that there are very small differences between ST and TGC density forecasts. The plots for JAP-US and Bitcoin show larger differences. For the former the TCG seems to perform better for the lower range whilst the ST performs better for the upper. For Bitcoin, it is apparent that the ST performs better overall, although for both tails the TGC does it better. Finally, Table 7 reports the VaR backtesting results for the ST-TGARCH, which provide further evidence of its similar performance in most cases with respect to the TGC-TGARCH (see Tables 4 and 5). Note that for Oil and Bitcoin there are more violations and higher AQL values under the ST.

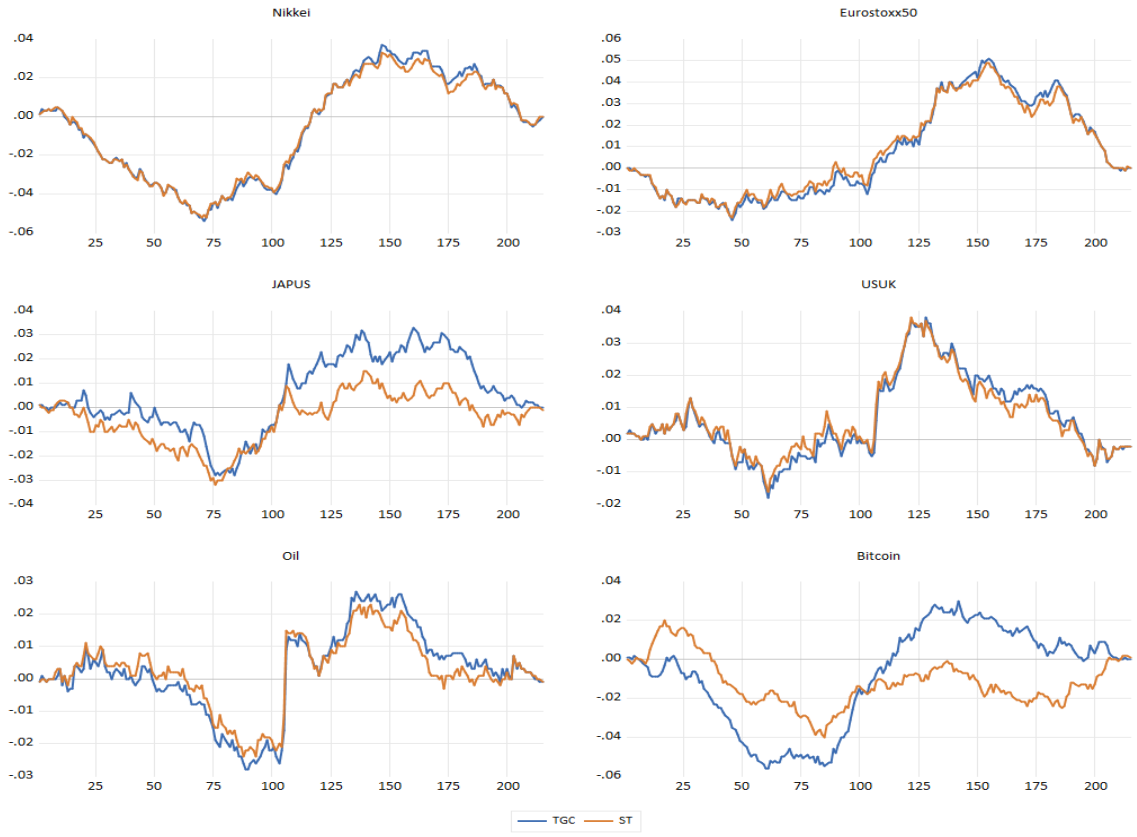


Figure 7: P-value discrepancy plots for PIT series obtained from TGC-TGARCH and ST-TGARCH models. Return series: Nikkei, Eurostoxx50, JAP-US, US-UK, Oil and Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns is from November 4, 2015 to July 31, 2018. Predictions: 1000.

[8] We thank the referee for pointing out this analysis of comparing distributions.

Table 6: Hansen's ST estimation results

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
υ	8.2869*	8.8975*	5.5675*	8.8001*	6.6174*	2.4338*
	(0.9007)	(1.1085)	(0.4258)	(1.2199)	(0.6128)	(0.1196)
δ	-0.0672*	-0.0683*	-0.0351	-0.0371*	-0.0758*	-0.0258
	(0.0185)	(0.0195)	(0.0179)	(0.0174)	(0.0179)	(0.0201)
LRV	-3.83*	-22.3*	2.48**	1.19	1.02	-0.17

This table presents the ML estimates of Hansen's ST parameters where $\delta \in (-1, 1)$ and $\upsilon > 2$ control for skewness and kurtosis, respectively. Mean and TGARCH equation estimates for ST are very similar to those under TGC (Table 2), thus they are not presented here to avoid duplicity. The sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from January 15, 1999 to January 14, 2019 (total: 5,217 obs.); and for Bitcoin returns it is from July 19, 2010 to July 31, 2018 (total: 2,935 obs.). Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*) and (**) indicate significance at 1% and 5% levels, respectively. LRV denote the Vuong's (1989) nonnested LR test statistic in (50) for $i = ST$ and $j = TGC$ under TGARCH model.

Table 7: Backtesting results for ST-TGARCH model

VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)	VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)
Nikkei				Eurostoxx50			
14	15.13	44	43.76	7	8.41	40^{c}	38.51^{u}
(0.0183)		(0.0734)		(0.0554)		(0.0930)	
JAP-US				US-UK			
13	13.93	46	47.27	11	14.91	55	54.61
(0.0060)		(0.0217)		(0.0432)		(0.0601)	
Oil				Bitcoin			
13	12.84	54^{c}	54.48^{c}	10	15.60	$67^{u,c}$	$62.10^{u,c}$
(0.0156)		(0.1360)		(0.2124)		(1.1522)	

This table shows the violations for VaR and the cumulative violations given in (61) for ES under the ST-TGARCH model. We also reports the significance for (i) the VaR backtesting tests in (58) and (59), and (ii) the ES backtesting tests in (62) and (63). We set $m = 5$ in the Box-Pierce test statistic for the two conditional backtests. The superscripts u and c indicate significance at five percent level for the unconditional and conditional backtests, respectively. The results of the AQL in (57) are provided in brackets. Return series: Nikkei, JAP-US, Oil, US-UK, Eurostoxx50, Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns is from November 4, 2015 to July 31, 2018. Predictions: 1000.

6 Extensions

There are at least two admissible extensions of the TGC framework in this paper. The first one follows recent results in the literature which show the importance of TV conditional higher-order moments with asymmetric response to positive and negative shocks; see e.g. Feunou, Jahan-Parvar and Tédongap (2016), and Lalancette and Simonato (2017). The TGC pdf allows a natural extension to model directly clustering and asymmetries in skewness and kurtosis. The time-varying version of the TGC pdf is denoted as TV-TGC. The second extension relates to the proposal of a mixture of TGC pdfs which can significantly improve the TGC ability to capture larger ranges of kurtosis.

6.1 Time-varying conditional higher-order moments

The conditional skewness and kurtosis of r_t are defined, respectively, as $s_{r,t} = E\left(\varepsilon_t^3 \,|I_{t-1}\right)/\sigma_t^3$ and $k_{r,t} = E\left(\varepsilon_t^4 \,|I_{t-1}\right)/\sigma_t^4$. If we let the pdf of x_t in (4) exhibit TV parameters and $\mu_t = \mu$ in (23), then $s_{r,t} = s_{z,t}$ and $k_{r,t} = k_{z,t}$ are now TV conditional daily higher-order moments where $s_{z,t}$ and $k_{z,t}$ are obtained by plugging $\theta_{i,t}$ into equations (10) and (11). The dynamics equation for $\theta_{i,t}$ is given by

$$\theta_{i,t} = \varphi_{0i} + \varphi_{1i}\theta_{i,t-1} + \varphi_{2i}^{+} z_{t-1}^{+} + \varphi_{2i}^{-} z_{t-1}^{-}. \tag{65}$$

This specification, although initially proposed by JR (2003) under Hansen's ST distribution for the innovation z_t, has also been implemented, among others, by Feunou et al. (2016) who consider the skewed generalized error (SGE) distribution; Anatolyev and Petukhov (2016) under the SGE assuming time variation only for the parameter more related to skewness; Lalancette and Simonato (2017) for the Johnson S_u distribution, and León and Ñíguez (2018) who develop the conditional SNP distribution in LMS (2009).

Table 8 only exhibits the ML parameter estimates for $\theta_{i,t}$ in (65). Several features can be observed. First, there is evidence that the TGARCH model under the TV-TGC does capture skewness and kurtosis clustering since φ_{11} and φ_{12} are significant for all return series.[9] Second, we find evidence of asymmetric response, of both skewness and kurtosis, to positive and negative shocks in all return series except Bitcoin. Third, the LR test (H_0: TGC and H_1: TV-TGC) and AIC show that the TV-TGC model provides statistically significant improvements of goodness-of-fit respecting the TGC. Fourth, it is also reported the mean and standard deviation values of the TV conditional (daily) higher-order moment series. These mean values are rather the same to those obtained under the TGC model with constant conditional higher-order moments. Figure 8 exhibits in panel 1 both the constant and TV daily conditional skewness series for all asset returns. The same graphics for the daily conditional kurtosis are exhibited in panel 2.

Finally, in order to provide more evidence on the significance of the returns' daily conditional skewness, we employ a distribution-based asymmetry measure defined as the probability difference between the return distribution upside potential and the downside loss.[10] This measure does provide more accurate information on the distribution asymmetry than the skewness, which limits to the third-moment. In particular, we use the conditional version of the excess tail probability (ETP) of Jiang, Wu, Zhou and Zhu (2020), which is

[9] The TGARCH parameter estimates are not included to save space, as they are very similar to those in Table 2.

[10] The LPM of order zero in (15), i.e. $m = 0$.

defined as

$$\begin{aligned} ETP_t(\tau) &= \int_{\tau}^{+\infty} f(r_t|I_{t-1})\,dr_t - \int_{-\infty}^{-\tau} f(r_t|I_{t-1})\,dr_t \\ &= 1 - Q\left(\frac{\tau-\kappa_{0t}}{\kappa_{1t}}|I_{t-1}\right) - Q\left(\frac{-\tau-\kappa_{0t}}{\kappa_{1t}}|I_{t-1}\right), \end{aligned} \tag{66}$$

where $f(\cdot|I_{t-1})$ is the conditional pdf for r_t in (43) with TV-TGC distribution for z_t, expressed in terms of the non-standardized TV-TGC pdf with cdf given by $Q(\cdot|I_{t-1})$ according to (8). For the threshold level τ, we finally set a value of two standard deviations in accordance to the returns' sample standard deviations in Table 1.[11] Hence, the first and second terms in (66) measure the cumulative chances of large gains and large losses, respectively. If ETP_t is positive (negative), the probability of large gains is higher (lower) than the probability of large (losses). Thus, we should expect a big coincidence between positive (negative) daily ETP and positive (negative) daily skewness values. Our results show a very high percentage of similarity between the two asymmetry measures. Namely, we find 98% for Nikkei, 95% for Eurostoxx, 96% for JAP-US, 97% for US-UK, 97% for Oil, and a little lower of 80% for Bitcoin.

In short, the TV-TGC model seems to gather well the dynamics in the higher-order moments of the asset returns. The forecasting performance of this model as well as confidence intervals based on the ETP measure should constitute further research steps as it falls out ot the scope of this paper.

6.2 Mixture of TGC distributions

In this section we analyze the MTGC distribution. Our aim is to develop a more flexible TGC density specification capable of providing higher ranges of kurtosis.

The random variable ε_t in (23) follows a two-component mixed Normal (M2N) with GARCH if the conditional distribution of ε_t is M2N with zero mean, i.e. $\varepsilon_t|I_{t-1} \sim M2N\left(\mathbf{p},\boldsymbol{\mu},\boldsymbol{\sigma}_t^2\right)$ where $\mathbf{p}=(p,1-p)'$ with $0<p<1$, $\boldsymbol{\mu}=(\mu_1,\mu_2)'$ and $\boldsymbol{\sigma}_t^2=\left(\sigma_{1t}^2,\sigma_{2t}^2\right)'$, and the conditional pdf is given by

$$h_{M2N}(\varepsilon_t|I_{t-1}) = \sum_{j=1}^{2} p_j \frac{1}{\sigma_{jt}}\phi\left(\frac{\varepsilon_t-\mu_j}{\sigma_{jt}}\right), \tag{67}$$

with $\mu_1=\left(1-p^{-1}\right)\mu_2$ in order to ensure that $E(\varepsilon_t|I_{t-1})=0$, and $\sigma_{jt}^2=\alpha_{0j}+\alpha_{1j}\varepsilon_{t-1}^2+\beta_j\sigma_{jt-1}^2$. More details of this model such as the conditions of stationarity, persistence and other properties can be seen in Haas, Mittnik and Paolella (2004) and Alexander and Lazar (2006).

Now, we consider that ε_t follows a two-component mixed TGC (M2TGC) with GARCH if the conditional distribution of ε_t is M2TGC with zero mean. Thus, $\varepsilon_t|I_{t-1} \sim M2TGC\left(\mathbf{p},\boldsymbol{\mu},\boldsymbol{\sigma}_t^2,\boldsymbol{\theta}_1,\boldsymbol{\theta}_2\right)$ such that $\mathbf{p}=(p,1-p)'$ with $0<p<1$, $\boldsymbol{\theta}_j=(\theta_{1j},\theta_{2j})'$, $\boldsymbol{\mu}$ is the location parameter vector and $\boldsymbol{\sigma}_t^2$ is the vector of strictly positive scale parameters, and the conditional pdf is given by

$$h_{M2TGC}(\varepsilon_t|I_{t-1}) = \sum_{j=1}^{2} p_j \frac{1}{b_j\sigma_{jt}} q_j(\varepsilon_{jt},\boldsymbol{\theta}_j), \tag{68}$$

where $\varepsilon_{jt}=\left(\varepsilon_t-\mu_j-a_j\sigma_{jt}\right)/\left(b_j\sigma_{jt}\right)$ with $\mu_1=\left(1-p^{-1}\right)\mu_2$ as the condition of zero mean, σ_{jt}^2 is a GARCH structure, and $q_j(\cdot)$ denotes the TGC pdf in (4) with $\boldsymbol{\theta}_j=(\theta_{1j},\theta_{2j})'$. Note that the pdf in (67) is nested in (68) when the parameter vector $\boldsymbol{\theta}_j=\mathbf{0}$ for $j=1,2$. The random variable ε_t can be expressed as

$$\varepsilon_t = \begin{cases} \mu_1+\sigma_{1t}z_{1t}, & \text{with prob. } p, \\ \mu_2+\sigma_{2t}z_{2t}, & \text{with prob. } 1-p, \end{cases} \tag{69}$$

[11] Jian et al. (2020) use a range from 1 to 1.5 standard deviations in their empirical analysis.

such that z_{jt} in (69) is the standardized TGC defined as $z_{jt} = a_j + b_j x_{jt}$ such that $x_{jt} \sim i.i.d.\ TGC(\boldsymbol{\theta}_j)$ with $q_j(\cdot)$ in (4) as pdf, $a_j = -b_j E_{q_j}[x_{jt}]$, $b_j = \left(E_{q_j}\left[x_{jt}^2\right] - E_{q_j}[x_{jt}]^2\right)^{-1/2}$ and the non-central moments: $E_{q_j}[x_{jt}] = 4\lambda_j\gamma_{1j}\gamma_{2j}$, $E_{q_j}\left[x_{jt}^2\right] = 6\lambda_j\gamma_{1j}^2 + 8\lambda_j\gamma_{2j}^2 + 1$ with $\gamma_{1j} = \theta_{1j}/\sqrt{3!}$ and $\gamma_{2j} = \theta_{2j}/\sqrt{4!}$, see Proposition 2. The conditional moments of ε_t, such that $E(\varepsilon_t | I_{t-1}) = 0$, are easily obtained by using the binomial expansion. Thus,

$$E\left(\varepsilon_t^k | I_{t-1}\right) = \sum_{j=1}^{2} p_j \sum_{i=0}^{k} \binom{k}{i} \mu_j^{k-i} \sigma_{jt}^i E_{q_j}[(a_j + b_j x_{jt})^i].$$

A more detailed study such as fitting this model to data falls out of the scope of the paper but it remains as an interesting avenue for further research.

7 Conclusions

We analyze the parametric properties of a new density obtained as a result of applying the Gallant and Nychka (1987) method to the Gram-Charlier (GC) density in Jondeau and Rockinger (2001) in order to ensure positivity within the whole parametric space. This density is referred to as the transformed GC (TGC). We provide a thorough analysis of the TGC's statistical properties deriving the conditions for unimodality, allowable ranges of skewness and kurtosis, closed-form expressions of the distribution function, moments, one-sided truncated moments, expected shortfall and lower partial moments. We also obtain the analytic expressions for the k-th order stationarity conditions and study the power-law tail property for the unconditional moments of the errors from the TGARCH model under the TGC density, or TGC-TGARCH model.

We also study the performance of the TGC-TGARCH model through an empirical application to forecasting financial risk. In order to isolate the effect of skewness, we compare the TGC with the symmetric GC of Zoia et al. (2018). The Normal distribution is used as the benchmark model. The model ability to forecast both cdf and pdf is tested through p-value discrepancy plots and strictly proper scoring rules in Amisano and Giacomini (2007), respectively. We also employ the VaR and ES backtesting methods of Du and Escanciano (2017). For comparative analysis purposes, we extend our empirical study to the TGARCH under the Hansen's skewed t distribution, or ST-TGARCH model. Our last section deals with two possible lines for further research with some initial results already presented here. Specifically, the first consists on an extension of the TGC model to the TGC with time-varying skewness and kurtosis; and the second the TGC mixture distribution as a way to augment the range of kurtosis provided by the TGC.

Finally, we also consider the following avenues as interesting for future research. First, the tail-index estimation through Kesten's equation for asymmetric GARCH-family models as an extension to the work of Zhang et al. (2019). Second, the multivariate extensions of the univariate distributions dealt within the paper. For instance, either GC or TGC distributions as possible marginal ones with dependence structures based on alternative copulas such as (i) the Fourier expansions approach of Ibragimov and Prokhorov (2017), or (ii) the dynamic copula extension of the multivariate skewed t distribution of Demarta and McNeil (2005), developed by Christoffersen, Errunza, Jacobs and Langlois (2012). Third, in the spirit of the latter research line, possible extensions of the multivariate SNP framework by Ñíguez and Perote (2016).

Table 8: TV-TGC model estimation results

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
φ_{01}	-0.1114*	-0.0178	-0.0157***	-0.0606*	-0.1568*	0.0427
	(0.050)	(0.028)	(0.021)	(0.031)	(0.034)	(0.032)
φ_{11}	-0.4894*	0.4464*	0.5002*	-0.5536*	-0.8741*	0.6635*
	(0.528)	(0.779)	(0.140)	(0.149)	(0.055)	(0.465)
φ_{21}^{+}	0.0473	0.0125	0.0422**	0.0582***	0.0262	-0.0292
	(0.029)	(0.033)	(0.030)	(0.029)	(0.016)	(0.073)
φ_{21}^{-}	0.0152	0.0339	0.0431	0.0113	0.0001	0.1195
	(0.026)	(0.043)	(0.021)	(0.023)	(0.014)	(0.066)
φ_{02}	0.4853	0.0197	0.5066*	0.5623*	0.0751	1.5115*
	(0.115)	(0.036)	(0.131)	(0.075)	(0.049)	(0.156)
φ_{12}	-0.7134*	0.7005*	-0.3372*	-0.9621*	0.6044*	-0.4858*
	(0.221)	(0.092)	(0.235)	(0.011)	(0.109)	(0.140)
φ_{22}^{+}	0.0778	0.1978*	0.1222	-0.0246	0.1593*	-0.2218
	(0.057)	(0.052)	(0.052)	(0.018)	(0.050)	(0.082)
φ_{22}^{-}	-0.0977***	-0.0006	-0.1574*	0.0194	-0.0466	-0.0626
	(0.041)	(0.038)	(0.051)	(0.017)	(0.039)	(0.125)
$LL_{TV\text{-}TGC}$	-8731.26	-8643.46	-4941.91	-4045.72	-11226.51	-4607.26
LR	13.74*	23.01*	21.47*	7.59*	17.32*	16.64*
$AIC_{TV\text{-}TGC}$	3.3497	3.3161	1.8970	1.5535	4.3063	3.1440
$skew_{TGC}$	-0.1803	-0.1463	-0.1157	-0.0787	-0.2308	-0.0687
$\overline{skew}_{TV\text{-}TGC}$	-0.1821	-0.1236	-0.0954	-0.0756	-0.2213	-0.0169
s.d. $skew_{TV\text{-}TGC}$	(0.105)	(0.066)	(0.144)	(0.119)	(0.088)	(0.241)
$kurt_{TGC}$	3.8682	3.7845	4.3499	3.7736	4.0881	5.5585
$\overline{kurt}_{TV\text{-}TGC}$	3.9016	3.8786	4.3078	3.7551	4.0909	5.4742
s.d. $kurt_{TV\text{-}TGC}$	(0.262)	(0.441)	(0.289)	(0.156)	(0.307)	(0.278)

The parameter dynamics implied in the TV-TGC distribution of z_t: $\theta_{i,t} = \varphi_{0i} + \varphi_{1i}\theta_{i,t-1} + \varphi_{2i}^{+} z_{t-1}^{+} + \varphi_{2i}^{-} z_{t-1}^{-},\ i = 1, 2.$ This table presents ML estimates of the TV-TGC coefficients in the previous parameter equations of $\theta_{i,t}$ for percent log return series. The sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from January 15, 1999 to January 14, 2019 (total: 5,217 obs.); and for Bitcoin returns it is from July 19, 2010 to July 31, 2018 (total: 2,935 obs.). Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*), (**) and (***) indicate significance at 1%, 5% and 10% levels, respectively. $LL_{TV\text{-}TGC}$ denotes the log-likelihood value for the TV-TGC. LR denotes the likelihood ratio test statistic for TV-TGC versus TGC. Finally, $AIC_{TV\text{-}TGC}$ denotes the Akaike information criterion (the lower the AIC value, the better the goodness-of-fit). Both $skew_{TGC}$ and $kurt_{TGC}$ correspond to the skewness and kurtosis for the standardized return z_t under the TGC model with constant conditional skewness and kurtosis. Both $\overline{skew}_{TV\text{-}TGC}$ and $\overline{kurt}_{TV\text{-}TGC}$ correspond to the sample means of the daily time-varying conditional skewness and kurtosis series. Finally, the standard deviations for the skewness and kurtosis series are in parentheses below $\overline{skew}_{TV\text{-}TGC}$ and $\overline{kurt}_{TV\text{-}TGC}$.

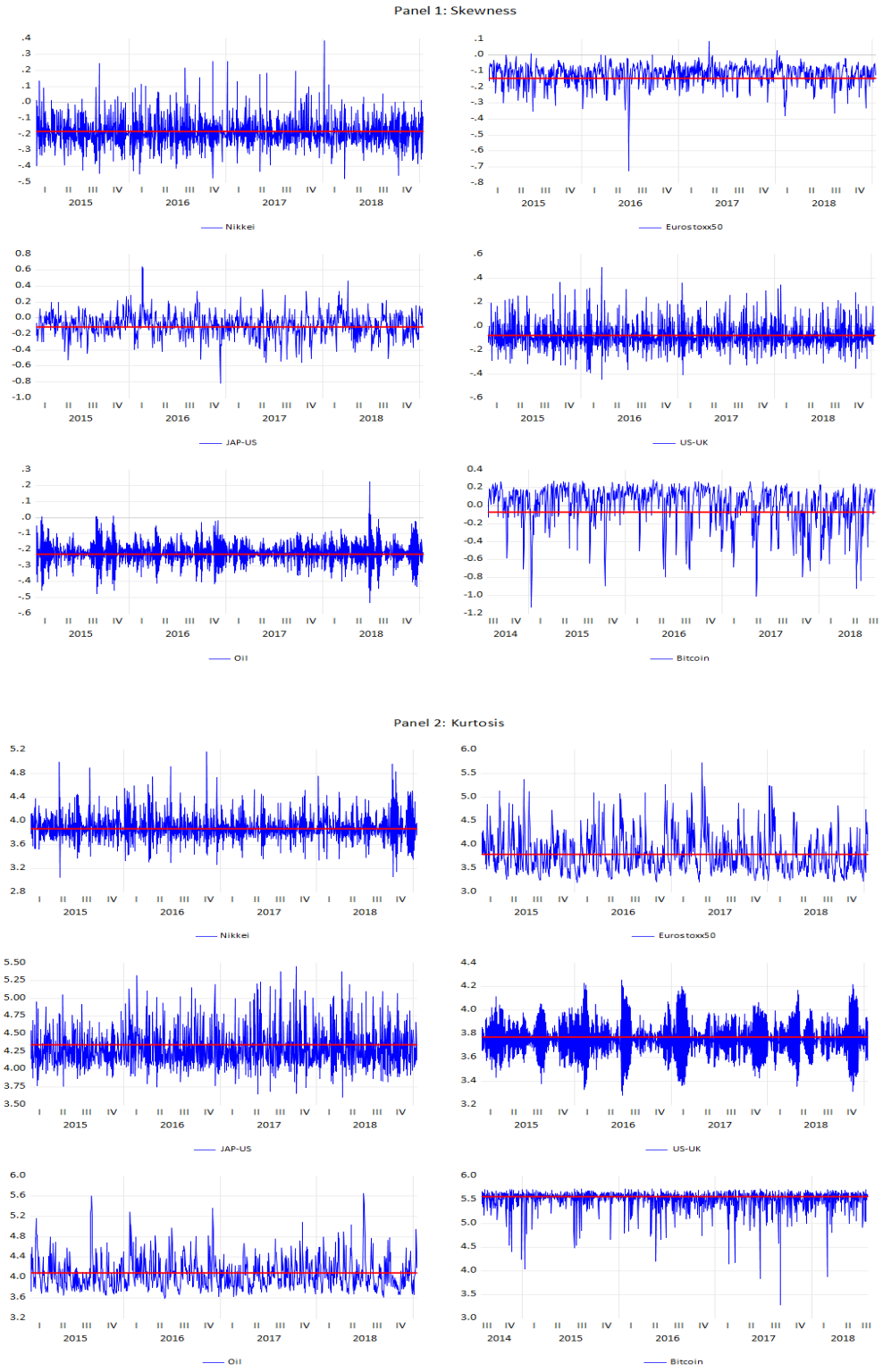


Figure 8: Plots of series of TV-TGC conditional skewness and kurtosis. The horizontal line correspond to the estimated skewness and kurtosis from the TGC model. The sample covers the last 1,000 observations of the whole sample period, for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin returns it is from April 11, 2015 to July 31, 2018.

Appendix 1 (Hermite polynomials)

The normalized Hermite polynomials, $H_k(x)$, can be expressed recursively for $k \geq 2$ as

$$H_k(x) = \frac{xH_{k-1}(x) - \sqrt{k-1}H_{k-2}(x)}{\sqrt{k}}, \tag{70}$$

with initial conditions $H_0(x) = 1$ and $H_1(x) = x$. The set $\{H_k(x)\}_{k \in N}$ constitutes an orthonormal basis with regard to a weighting function $\phi(x)$, which is the pdf of the standard Normal distribution. The orthonormality property means that $E_\phi[H_k(x)H_l(x)] = \delta_{kl}$, where δ_{kl} is the Kronecker delta ($\delta_{kl} = 1$ if $k = l$ while $\delta_{kl} = 0$ otherwise) and the operator $E_\phi[\cdot]$ takes the expectation with $\phi(x)$ as pdf. It is verified that even and odd degree Hermite polynomials lead to even and odd functions, respectively. We can also express $H_k(x)$ in terms of x^k by considering the following result from Blinnikov and Moessner (1998):

$$H_k(x) = \sqrt{k!} \sum_{n=0}^{[k/2]} \frac{(-1)^n}{n!\,(k-2n)!2^n} x^{k-2n}, \tag{71}$$

where $[\cdot]$ rounds its argument to the nearest integer toward zero. Given (71), we can now write x^k in terms of $H_k(x)$ as follows:

$$x^k = \sqrt{k!}H_k(x) - \sum_{n=1}^{[k/2]} \frac{(-1)^n\, k!}{n!\,(k-2n)!2^n} x^{k-2n}. \tag{72}$$

If we substitute recursively the powers of x on the right side of equation (72) by the same equation, we finally obtain x^k as a polynomial transformation of a set of Hermite polynomials of degrees lower or equal than k.

Appendix 2 (Proofs)

Let $m_k(x)$ denote the one-sided truncated standard Normal moment of order k given by $E\left[z^k \,|z \leq x\right]$ with pdf $\phi(z)/\Phi(x)$, then a recursive formula (see Liquet and Nazarathy, 2015) is obtained for $k \geq 1$ as $m_k(x) = (k-1)\, m_{k-2}(x) - x^{k-1}\phi(x)/\Phi(x)$, where $m_{-1}(x) = 0$ and $m_0(x) = 1$. Let $B_k(x) = \int_{-\infty}^{x} z^k \phi(z) dz$, then $B_k(x) = m_k(x)\,\Phi(x)$. We can obtain a recursive expression for $B_k(x)$ as

$$B_k(x) = (k-1)\, B_{k-2}(x) - x^{k-1}\phi(x), \qquad k \geq 2 \tag{73}$$

where $B_0(x) = \Phi(x)$ and $B_1(x) = -\phi(x)$. The recursion formula in (73) is also obtained, in a very slightly different way, in Skoulakis (2019) and it will be useful for all our proofs.

a) Equations of $\Gamma_{kij}(\cdot)$. Consider $\xi_{kij} = 1/\sqrt{k!i!j!}$, the equation in (71) and $B_k(x)$ in (73), then we obtain all equations $\Gamma_{kij}(\cdot)$ in Proposition 1 that we need to compute (8), (13), (17), (32) and Corollary 5. Then, we have

i) For $k=0$:

$$\begin{aligned}
&\Gamma_{030}(x)=\xi_{030}\left[B_3(x)-3B_1(x)\right],\\
&\Gamma_{040}(x)=\xi_{040}\left[B_4(x)-6B_2(x)+3B_0(x)\right],\\
&\Gamma_{033}(x)=\xi_{033}\left[B_6(x)-6B_4(x)+9B_2(x)\right],\\
&\Gamma_{034}(x)=\xi_{034}[B_7(x)-9B_5(x)+21B_3(x)-9B_1(x)],\\
&\Gamma_{044}(x)=\xi_{044}\left[B_8(x)-12B_6(x)+42B_4(x)-36B_2(x)+9B_0(x)\right].
\end{aligned} \tag{74}$$

ii) For $k=1$:

$$\begin{aligned}
&\Gamma_{100}(x)=B_1(x),\\
&\Gamma_{130}(x)=\xi_{130}\left[B_4(x)-3B_2(x)\right],\\
&\Gamma_{140}(x)=\xi_{140}\left[B_5(x)-6B_3(x)+3B_1(x)\right],\\
&\Gamma_{133}(x)=\xi_{133}\left[B_7(x)-6B_5(x)+9B_3(x)\right],\\
&\Gamma_{134}(x)=\xi_{134}\left[B_8(x)-9B_6(x)+21B_4(x)-9B_2(x)\right],\\
&\Gamma_{144}(x)=\xi_{144}\left[B_9(x)-12B_7(x)+42B_5(x)-36B_3(x)+9B_1(x)\right].
\end{aligned} \tag{75}$$

iii) For $k=2$:

$$\begin{aligned}
&\Gamma_{200}(x)=\xi_{200}\left[B_2(x)-B_0(x)\right],\\
&\Gamma_{230}(x)=\xi_{230}\left[B_5(x)-4B_3(x)+3B_1(x)\right],\\
&\Gamma_{240}(x)=\xi_{240}\left[B_6(x)-7B_4(x)+9B_2(x)-3B_0(x)\right],\\
&\Gamma_{233}(x)=\xi_{233}\left[B_8(x)-7B_6(x)+15B_4(x)-9B_2(x)\right],\\
&\Gamma_{234}(x)=\xi_{234}\left[B_9(x)-10B_7(x)+30B_5(x)-30B_3(x)+9B_1(x)\right],\\
&\Gamma_{244}(x)=\xi_{244}\left[B_{10}(x)-13B_8(x)+54B_6(x)-78B_4(x)+45B_2(x)-9B_0(x)\right].
\end{aligned} \tag{76}$$

iv) For $k=3,4$:

$$\begin{aligned}
&\Gamma_{333}(x)=\xi_{333}\left[B_9(x)-9B_7(x)+27B_5(x)-27B_3(x)\right],\\
&\Gamma_{334}(x)=\xi_{334}\left[B_{10}(x)-12B_8(x)+48B_6(x)-72B_4(x)+27B_2(x)\right],\\
&\Gamma_{344}(x)=\xi_{344}\left[B_{11}(x)-15B_9(x)+78B_7(x)-162B_5(x)+117B_3(x)-27B_1(x)\right],\\
&\Gamma_{444}(x)=\xi_{444}\left[B_{12}(x)-18B_{10}(x)+117B_8(x)-324B_6(x)+351B_4(x)-162B_2(x)+27B_0(x)\right].
\end{aligned} \tag{77}$$

b) Proof of Proposition 2 (moments). According to (72), we can express the first four powers of x as a function of the Hermite polynomials, i.e. $x=H_1(x)$, $x^2=\sqrt{2}H_2(x)+H_0(x)$, $x^3=\sqrt{3!}H_3(x)+3H_1(x)$ and $x^4=\sqrt{4!}H_4(x)+6\sqrt{2}H_2(x)+3H_0(x)$. Then, we can easily obtain $E_q\left[x^k\right]=\int_{-\infty}^{+\infty}x^k q(x)\,dx$ as follows:

$$\begin{aligned}
&E_q[x]=2\lambda\gamma_1\gamma_2 A_{134},\\
&E_q\left[x^2\right]=\sqrt{2}\lambda\gamma_1^2 A_{233}+\sqrt{2}\lambda\gamma_2^2 A_{244}+1,\\
&E_q\left[x^3\right]=2\sqrt{3!}\lambda\gamma_1+2\sqrt{3!}\lambda\gamma_1\gamma_2 A_{334}+6\lambda\gamma_1\gamma_2 A_{134},\\
&E_q\left[x^4\right]=2\sqrt{4!}\lambda\gamma_2+\sqrt{4!}\lambda\gamma_1^2 A_{334}+\sqrt{4!}\lambda\gamma_2^2 A_{444}+6\sqrt{2}\lambda\gamma_1^2 A_{233}+6\sqrt{2}\lambda\gamma_2^2 A_{244}+3,
\end{aligned} \tag{78}$$

where $A_{ijk}=E_\phi\left[H_{ijk}(x)\right]=\int_{-\infty}^{+\infty}H_{ijk}(x)\,\phi(x)\,dx$ with $H_{ijk}(x)=H_i(x)\,H_j(x)\,H_k(x)$ verifying that

$A_{ijk} = 0$ if $i + j + k$ is an odd number and

$$\begin{aligned}
A_{233} &= \xi_{233}\left[n_8 - 7n_6 + 15n_4 - 9n_2\right], \\
A_{134} &= \xi_{134}\left[n_8 - 9n_6 + 21n_4 - 9n_2\right], \\
A_{244} &= \xi_{244}\left[n_{10} - 13n_8 + 54n_6 - 78n_4 + 45n_2 - 9\right], \\
A_{334} &= \xi_{334}\left[n_{10} - 12n_8 + 48n_6 - 72n_4 + 27n_2\right], \\
A_{444} &= \xi_{444}\left[n_{12} - 18n_{10} + 117n_8 - 324n_6 + 351n_4 - 162n_2 + 27\right],
\end{aligned} \tag{79}$$

where $\xi_{ijk} = 1/\sqrt{i!j!k!}$ and $n_{2k} = E_\phi\left[x^{2k}\right] = \frac{(2k)!}{2^k k!}$. Then, $A_{233} = 3\sqrt{2}$, $A_{134} = 2$, $A_{244} = 4\sqrt{2}$, $A_{334} = 3\sqrt{6}$, $A_{444} = 6\sqrt{6}$ and finally, we obtain (9).

c) Equations of ω_k under Normal-TGARCH. If $z_t \sim$ *i.i.d.* $N(0,1)$, then $E\left(z_t^-\right) = -1/\sqrt{2\pi}$, $E\left[\left(z_t^-\right)^2\right] = 1/2$, $E\left[\left(z_t^-\right)^3\right] = -\sqrt{2/\pi}$, $E\left[\left(z_t^-\right)^4\right] = 3/2$, and $\varpi_k = E\left(c_t^k\right)$ for $k = 1, 2, 4$ given by

$$\begin{aligned}
\varpi_1 &= \beta + \frac{1}{\sqrt{2\pi}}\left(\alpha_1^- + \alpha_1^+\right), \\
\varpi_2 &= \beta^2 + \sqrt{\frac{2}{\pi}}\beta\left(\alpha_1^- + \alpha_1^+\right) + \frac{1}{2}\left[\left(\alpha_1^-\right)^2 + \left(\alpha_1^+\right)^2\right], \\
\varpi_4 &= \beta^4 + 2\sqrt{\frac{2}{\pi}}\beta^3\left(\alpha_1^- + \alpha_1^+\right) + 3\beta^2\left[\left(\alpha_1^-\right)^2 + \left(\alpha_1^+\right)^2\right] + 4\sqrt{\frac{2}{\pi}}\beta\left[\left(\alpha_1^-\right)^3 + \left(\alpha_1^+\right)^3\right] + \frac{3}{2}\left[\left(\alpha_1^-\right)^4 + \left(\alpha_1^+\right)^4\right].
\end{aligned} \tag{80}$$

d) Equations of $E\left(c_t^k\right)$ under NAGARCH. Let $\varepsilon_t = \sigma_t z_t$ with $z_t \sim$ *i.i.d.* $D(0,1)$, then the NAGARCH model of Engle and Ng (1993) for the conditional variance is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1(\varepsilon_{t-1} - \delta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2, \tag{81}$$

with $\alpha_0 > 0$, $\beta \geq 0$, $\alpha_1 \geq 0$ and $\delta \in \mathbb{R}$. Following He and Terasvirta (1999), we can rewritte (81) as $\sigma_t^2 = \alpha_0 + c_{t-1}\sigma_{t-1}^2$ such that $c_t = \beta + \alpha_1(z_t - \delta)^2$. We can obtain recursively $\varpi_k = E\left[c_t^k\right]$ for $k \geq 2$:

$$\varpi_k = \alpha_1^k E\left[z_t^{2k}\right] + \alpha_1^k \sum_{j=1}^{2k} \binom{2k}{j} (-\delta)^j E\left[z_t^{2k-j}\right] - \sum_{j=1}^{k} \binom{k}{j} (-\beta)^j \varpi_{k-j}, \tag{82}$$

where

$$\varpi_1 = \beta + \alpha_1\left(1 + \delta^2\right). \tag{83}$$

Note that $\varpi_1 < 1$ and $\varpi_2 < 1$ are the second and fourth-order stationarity conditions such that $E\left(\sigma_t^2\right) < \infty$ and $E\left(\sigma_t^4\right) < \infty$, respectively. For $k = 2$ in (82), we obtain:

$$\varpi_2 = \beta^2 + 2\alpha_1\beta\left(1 + \delta^2\right) + \alpha_1^2(k_z - 4\delta s_z + 6\delta^2 + \delta^4), \tag{84}$$

where s_z is the skewness and k_z is the kurtosis of z_t.

Appendix 3 (TGC-NAGARCH model)

Table A3 displays the estimation results for the NAGARCH under the TGC pdf.

Table A3: TGC-NAGARCH estimation results

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
μ	-0.0123	0.0077	0.0028	0.0006	0.0131	0.1076*
	(0.017)	(0.016)	(0.009)	(0.007)	(0.027)	(0.071)
α_0	0.0469*	0.0319*	0.0050*	0.0026*	0.0270*	1.0026*
	(0.012)	(0.007)	(0.002)	(0.001)	(0.010)	(0.642)
β	0.8614*	0.8513*	0.9474*	0.9404*	0.9433*	0.7191*
	(0.020)	(0.017)	(0.012)	(0.015)	(0.011)	(0.060)
α_1	0.0814*	0.0718*	0.0429*	0.0521*	0.0401*	0.3998*
	(0.011)	(0.010)	(0.008)	(0.016)	(0.007)	(0.107)
δ	0.6995*	0.9632*	0.1191	0.1175	0.5929*	0.0680
	(0.101)	(0.135)	(0.145)	(0.130)	(0.118)	(0.093)
θ_1	-0.0677*	-0.0563*	-0.0404*	-0.0218	-0.0789*	-0.0626*
	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)	(0.029)
θ_2	0.2966*	0.2795*	0.4604*	0.2859*	0.3763*	0.9932*
	(0.034)	(0.035)	(0.035)	(0.037)	(0.035)	(0.053)
ϖ_1	0.9826	0.9898	0.9909	0.9933	0.9975	1.1208
σ_ε	1.6423	1.7669	0.7379	0.6267	3.2978	- - - -
ϖ_2	1.0005	1.0161	0.9882	0.9943	1.0031	1.9950
LRV	1.1560	1.1505	-1.0285	0.4147	0.8011	0.1631

Model specification: $r_t = \mu + \varepsilon_t,\ \ \varepsilon_t = \sigma_t z_t,\ \ \ z_t \sim i.i.d.\ TGC\left(0, 1, \boldsymbol{\theta}\right),\ \ \sigma_t^2 = \alpha_0 + \alpha_1(\varepsilon_{t-1} - \delta\sigma_{t-1})^2 + \beta\sigma_{t-1}^2$.

This table presents ML estimates of parameters of the TGC-NAGARCH model for percent log return series. The sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from January 15, 1999 to January 14, 2019 (total: 5,217 obs.); and for Bitcoin returns it is from July 19, 2010 to July 31, 2018 (total: 2,935 obs.). Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*) and (**) indicate significance at 1% and 5% levels, respectively. Let $\varpi_k = E\left(c_t^k\right)$ where $\varpi_1 = \beta + \alpha_1\left(1 + \delta^2\right)$, then the second-order stationarity condition of the NAGARCH must verify $\varpi_1 < 1$. The implied unconditional standard deviation of ε_t is given by $\sigma_\varepsilon = \sqrt{\alpha_0 / \left(1 - \varpi_1\right)}$. The fourth-order stationarity condition under TGC-NAGARCH must verify $\varpi_2 < 1$ with ϖ_2 given in (84). Finally, LRV is the nonnested Vuong's (1989) LR test statistic in (50) for $i = TGC$ and $j = GCK$ under NAGARCH.

Appendix 4 (Tail-index estimation)

Table A4 exhibits the robust estimates of ζ (instead of the popular Hill's estimator), denoted as $\widehat{\zeta}_{RS}$, obtained through the OLS approach using the log-log rank-size (RS) regressions from Gabaix and Ibragimov (2011). The standard error of $\widehat{\zeta}_{RS}$ is equal to $\widehat{\zeta}_{RS}\sqrt{2/n}$ where $n < T$ with T as the total number of observations. For n, we set the most commonly used values for extreme observations, i.e. $n = k_0 \times T$ with $k_0 = 5\%, 10\%$. The 95% confidence interval for ζ is given by $\widehat{\zeta}_{RS} \pm 1.96 \times \widehat{\zeta}_{RS}\sqrt{2/n}$.

Table A4: Tail-index estimates and confidence intervals

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
Panel 1: Truncation 10%						
$\widehat{\zeta}_{RS}$	3.42	3.15	3.33	3.49	3.23	2.52
$s.e.\,(\widehat{\zeta}_{RS})$	0.21	0.19	0.21	0.22	0.20	0.21
$95\%\,CI_{RS}$	(3.00, 3.83)	(2.77, 3.53)	(2.93, 3.74)	(3.06, 3.91)	(2.84, 3.62)	(2.11, 2.93)
Panel 2: Truncation 5%						
$\widehat{\zeta}_{RS}$	3.64	3.55	3.65	3.41	3.51	2.91
$s.e.\,(\widehat{\zeta}_{RS})$	0.32	0.31	0.32	0.30	0.31	0.34
$95\%\,CI_{RS}$	(3.02, 4.28)	(2.94,4.15)	(3.02, 4.28)	(2.83, 4.00)	(2.91, 4.11)	(2.24, 3.57)

$\widehat{\zeta}_{RS}$ denotes the tail-index estimate, the standard error is $s.e.\,(\widehat{\zeta}_{RS})$ and $95\%\,CI_{RS}$ denotes the confidence interval at 95% level for the tail index ζ. The sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from January 15, 1999 to January 14, 2019 ($T = 5,217$ obs.); and for Bitcoin returns it is from July 19, 2010 to July 31, 2018 ($T = 2,935$ obs.).

Appendix 5 (Implied tail index)

We show in Table A5 the results of the simulation procedure to compute numerically the positive real number $\zeta > 0$ from Kesten's equation: $E\left(c_t^{\zeta}\right) = 1$ with $c_t = \beta + \alpha_1(z_t - \delta)^2$ under the NAGARCH, see section d) in Appendix 2, with $\zeta^* = 2\zeta$ as the corresponding tail index. However, respecting the TGARCH the tail index is directly the solution ζ obtained from Kesten's equation with c_t in (26). For simplicity in the simulation analysis, without loss of generality, we consider the NAGARCH under the GC pdf in (1) with parameters: $sk = 0$ and alternative values of ek. Let $\varpi_k = E\left(c_t^k\right)$, then the corresponding ϖ_2 values by only changing ek are obtained according to (84). We fix the value of ϖ_1 in (83) to 0.9957. This value is obtained according to the QML NAGARCH parameter estimates for our Oil return series, i.e. $\widehat{\alpha}_0 = 0.0325$, $\widehat{\alpha}_1 = 0.0421$, $\widehat{\beta} = 0.9399$ and $\widehat{\delta} = 0.5703$. We find, as expected, that the higher ϖ_2 the lower ζ^*.

To obtain ζ from Kesten's equation, we adopt the following procedure: *(i)* we draw a random sample for $z_t \sim GC\,(0, ek)$ of size $T_s = 500,000$ by using the GC cdf in (3), *(ii)* take the sample average of $\left\{c_i^{\zeta_0}\right\}_{i=1}^{T_s}$, with $c_i = \widehat{\beta} + \widehat{\alpha}_1(z_i - \widehat{\delta})^2$ such that $\{z_i\}_{i=1}^{T_s}$ is the GC random sample, as the estimation of $E\left(c_t^{\zeta_0}\right)$ with ζ_0 as a starting value of the unknown ζ, *(iii)* solve the nonlinear $E\left(c_t^{\zeta}\right) = 1$ using the 'fsolve' function in Matlab, and *(iv)* the tail index is computed as $\zeta^* = 2\zeta$. Our simulation procedure is rather similar to Chan, Peng and Zhang (2012). Indeed, we only use one draw of a large sample size to compute ζ, whilst the former authors employ an average value of ζ estimates from 1,000 random samples of much lower size.

We have checked that our procedure does approximate considerably well the ζ true value. For instance, if we set $ek = 1$, then the solution of Kesten's equation is: $\zeta = 2.1729$ (i.e. $\zeta^* = 4.35$). Besides, our simulated averages, as an approximation to ϖ_k for different k values, are: $\widehat{\varpi}_1 = 0.9955$ (true value is $\varpi_1 = 0.9957$), $\widehat{\varpi}_2 = 0.9986$ (true value is $\varpi_2 = 0.9990$), $\widehat{\varpi}_3 = 1.0114$, $\widehat{\varpi}_4 = 1.0375$, $\widehat{\varpi}_5 = 1.0822$, etc. Thus, the higher k the higher ϖ_k. As a remark, we find the unconditional moments are finite up to order four. In conclusion, as exhibited in Table A5, for $ek \geq 2$ ($ek \leq 1$) the unconditional moments exist only up to the third (fourth) order.

Table A5: Implied tail index

ek	0.5	1	2	3	4
ϖ_2	0.9981	0.9990	1.0008	1.0025	1.0043
ζ^*	4.62	4.35	3.95	3.68	3.47

Acknowledgements

We are grateful to participants at the 44th Symposium of the Spanish Economic Association (Alicante, 2019) for helpful comments, discussions and suggestions. Of course, the usual caveat applies. Financial support from the Spanish Ministry of Economy and Competitiveness through grant ECO2017-87069-P is gratefully acknowledged by the first author.

References

[1] Alexander, C., Lazar, E., 2006. Normal mixture GARCH (1, 1): Applications to exchange rate modelling. *Journal of Applied Econometrics* 21(3), 307-336.

[2] Alizadeh, A.H., Gabrielsen, A., 2013. Dynamics of credit spread moments of European corporate bond indexes. *Journal of Banking & Finance* 37(8), 3125-3144.

[3] Amédée-Manesme, C.O., Barthélémy, F., Maillard., D. 2019. Computation of the corrected Cornish–Fisher expansion using the response surface methodology: application to VaR and CVaR. *Annals of Operations Research* 281, 423-453.

[4] Amisano, G., Giacomini, R. 2007. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics* 25, 177-190.

[5] Anatolyev, S., Petukhov, A., 2016. Uncovering the skewness news impact curve. *Journal of Financial Econometrics* 14(4), 746-771.

[6] Auer, B.R., 2015. Superstitious seasonality in precious metals markets? Evidence from GARCH models with time-varying skewness and kurtosis. *Applied Economics* 47(27), 2844-2859.

[7] 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.

[8] Beber, A., Brandt, M.W., 2006. The effect of macroeconomic news on beliefs and preferences: Evidence from the options market. *Journal of Monetary Economics* 53(8), 1997-2039.

[9] Blinnikov, S., Moessner, R., 1998. Expansions for nearly Gaussian distributions. *Astronomy and Astrophysics Supplement Series* 130, 193-205.

[10] Chan, N.H., Peng, L., Zhang, R., 2012. Interval estimation of the tail index of a GARCH (1, 1) model. *Test*, 21(3), 546-565.

[11] Chan, N.H., Li, D., Peng, L., Zhang, R., 2013. Tail index of an AR (1) model with ARCH (1) errors. *Econometric Theory* 29, 920-940.

[12] Cheng, X., Li, W.K., Philip, L.H., Zhou, X., Wang, C., Lo, P.H., 2011. Modeling threshold conditional heteroscedasticity with regime-dependent skewness and kurtosis. *Computational Statistics & Data Analysis* 55(9), 2590-2604.

[13] Christoffersen, P.F. 2012. *Elements of Financial Risk Management*, 2nd edn. Academic Press, New York.

[14] Christoffersen, P., Errunza, V., Jacobs, K., Langlois, H. 2012. Is the Potential for International Diversification Disappearing? A Dynamic Copula Approach. *Review of Financial Studies* 25, 3711-3751.

[15] Corrado, C., 2007. The hidden martingale restriction in Gram-Charlier option prices. *Journal of Futures Markets* 27(6), 517-534.

[16] 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.

[17] Davidson, R., MacKinnon, J.G., 1998. Graphical Methods for Investigating the Size and Power of Hypothesis Tests. *The Manchester School* 66, 1-26.

[18] Del Brio, E.B., Mora-Valencia, A., Perote, J., 2020. Risk quantification for commodity ETFs: Backtesting value-at-risk and expected shortfall. *International Review of Financial Analysis* 70, 101163.

[19] Demarta, S., McNeil, A.J., 2005. The t copula and related copulas. *International Statistical Review* 73(1), 111-129.

[20] Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13(3), 253-263.

[21] Diebold, F.X., Gunther, T.A., Tay, S.A., 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39 (4), 863-883.

[22] Draper, N., Tierney, D., 1972. Regions of positive and unimodal series expansion of the Edgeworth and Gram-Charlier approximation. *Biometrika* 59, 463-465.

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

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

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

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

[27] Fishburn, P.C., 1977. Mean-risk analysis with risk associated with below.target returns. *American Economic Review* 67(2), 116-126.

[28] Francq, C. and Zakoïan, J.M. 2010. *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.

[29] Gabaix, X., 2009. Power laws in economics and finance. *Annual Review of Economics* 1(1), 255-294.

[30] Gabaix, X., Ibragimov, R., 2011. Rank-1/2: a simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics* 29, 24-39.

[31] Gabrielsen, A., Kirchner, A., Liu, Z., Zagaglia, P., 2012. Forecasting value-at-risk with time-varying variance, skewness and kurtosis in an exponential weighted moving average framework. *Annals of Financial Economics* 10(1), 1-29.

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

[33] Gallant, A.R., Tauchen, G., 1989. Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing implications. *Econometrica* 57, 1091-1120.

[34] Haas, M., Mittnik, S. and Paolella, M.S., 2004. Mixed normal conditional heteroskedasticity. *Journal of Financial Econometrics* 2(2), 211-250.

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

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

[37] Ibragimov, M., Ibragimov, R., Walden, J. 2015. *Heavy-tailed distributions and robustness in economics and finance*. Springer.

[38] Ibragimov, R., Prokhorov, A., 2017. *Heavy tails and copulas: topics in dependence modelling in economics and finance*. World Scientific.

[39] Jiang, L., Wu, K., Zhou, G., Zhu, Y., 2020. Stock return asymmetry: beyond skewness. *Journal of Financial and Quantitative Analysis* 55(2), 357-386.

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

[41] Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness and kurtosis: Existence, persistence, and comovements. *Journal of Economic Dynamics and Control* 27, 1699-737.

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

[43] Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica* 131, 207-248.

[44] Kräussl, R., Lehnert, T., Senulyté, S., 2016. Euro crash risk. *Journal of Empirical Finance* 38, 417-428.

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

[46] Lalancette, S., Simonato, J.G., 2017. The role of the conditional skewness and kurtosis in VIX index valuation. *European Financial Management* 23(2), 325-354.

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

[48] León, A., Rubio, G., Serna, G., 2005. Autoregressive conditional volatility, skewness and kurtosis. *Quarterly Review of Economics and Finance* 45, 599-618.

[49] 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.

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

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

[52] Liquet, B., Nazarathy, Y., 2015. A dynamic view to moment matching of truncated distributions. *Statistics and Probability Letters* 104, 87-93.

[53] Liu, X., Luger, R., 2015. Unfolded GARCH models. *Journal of Economic Dynamics and Control* 58, 186-217.

[54] Lönnbark, C., 2016. Approximation methods for multiple period Value at Risk and Expected Shortfall prediction. *Quantitative Finance* 16(6), 947-968.

[55] Mikosch, T., Starica, C., 2000. Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process. *The Annals of Statistics* 28 (5), 1427-1451.

[56] Narayan, P.K., Liu, R., 2018. A new GARCH model with higher moments for stock return predictability. *Journal of International Financial Markets, Institutions and Money* 56, 93-103.

[57] Ñíguez, T.M., Perote, J., 2012. Forecasting heavy-tailed densities with positive Edgeworth and Gram-Charlier expansions. *Oxford Bulletin of Economics and Statistics* 74(4), 600-627.

[58] Ñíguez, T.M., Perote, J. 2016. Multivariate moments expansion density: application of the dynamic equicorrelation model. *Journal of Banking and Finance* 72, S216-S232.

[59] Polanski, A., Stoja, E., 2010. Incorporating higher moments into value-at-risk forecasting. *Journal of Forecasting* 29(6), 523-535.

[60] 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.

[61] Skoulakis, G. 2019. Simulating from polynomial-normal distributions. *Communications in Statistics-Simulation and Computation* 48(2), 472-477.

[62] Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica,* 307-333.

[63] White, H., Kim, T.H., Manganelli, S., 2010. Modeling autoregressive conditional skewness and kurtosis with multi-quantile caviar. In *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle*, Oxford University Press.

[64] Wu, X., Xia, M., Zhang, H., 2020. Forecasting VaR using realized EGARCH model with skewness and kurtosis. *Finance Research Letters* 32, 101090.

[65] Zakoïan, J.M., 1994. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* 18 (5), 931-955.

[66] Zhang, R., Li, C., Peng, L., 2019. Inference for the tail index of a GARCH (1, 1) model and an AR (1) model with ARCH (1) errors. *Econometric Reviews* 38(2), 151-169.

[67] Zhu, D., 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.

[68] Zoia, M.G., 2010. Tailoring the gaussian law for excess kurtosis and skewness by hermite polynomials. *Communications in Statistics-Theory and Methods* 39, 52-64.

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