

WORKING PAPER SERIES

2020/001

Modeling asset returns under transformed Gram-Charlier

Ángel León

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

Trino-Manuel Níguez

Westminster Business School, University of Westminster

All rights reserved

Modeling asset returns under transformed Gram-Charlier

Ángel León

Dept. Fundamentos del Análisis Económico (FAE), Universidad de Alicante, Alicante 03690, Spain.

E-mail: aleon@ua.es

Trino-Manuel Níguez*

Westminster Business School, University of Westminster, London NW1 5UL, UK.

E-mail: t.m.niguez@westminster.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 transformed (TGC) density according to the Gallant and Nychka (1987) methodology. We derive TGC's parametric properties such as unimodality, cumulative distribution, higher-order moments and obtain closed-form formulae for expected shortfall (ES) and lower partial moments. In an empirical application, we backtest the density, Value-at-Risk and ES of several asset returns and show that our TGC provides accurate forecast for lower coverage levels. Finally, we present a TGC density with time-varying conditional skewness and kurtosis.

Keywords: Backtesting; Expected shortfall; Kurtosis; Skewness, Unimodality

JEL classification codes: C22, G11, G17.

*Corresponding author

1 Introduction

Densities based on polynomial expansions (PE) have attracted great attention to model the departures from normality shown in the empirical return distributions. Working with PE densities, we select a parent distribution whose third and fourth moments can be modified to match the empirical distribution by expanding a polynomial over the parent distribution. Bagnato, Poti and Zoia (2015) present a simple theorem that links higher-order moments (skewness and kurtosis) of both distributions. They obtain different orthogonal polynomials according to the parent density. The Gram-Charlier (GC) in Jondeau-Rockinger (2001) (henceforth, JR) is the PE density with the standard normal as parent density. It has become very appealing since the two parameters implied in the GC density directly correspond to the skewness and excess kurtosis. See, among others, Beber and Brandt (2006), Christoffersen and Diebold (2006), Polanski and Stoja (2010), Cheng, Philip, Zhou, Wang and Lo (2011), Níguez and Perote (2012), Liu and Luger (2015), Lönnbark (2016), Del Brio, Mora-Valencia and Perote (2017), León and Moreno (2017) and Zoia, Biffi and Nicolussi (2018). A special mention deserves successful application of GC densities to option pricing such as Corrado and Su (1996), Corrado (2007), León Mencía and Sentana (2009) and Schlögl (2013) and references therein.

An obvious problem of the GC density is that their unconstrained parameters can render negative probabilities. This issue has mainly been dealt with two approaches. First, imposing parameter restrictions that yield positive GC densities according to JR (2001). Second, transformations based on the methodology of Gallant and Nychka (1987), (henceforth, GN). Indeed, the GN approach is followed by León, Rubio and Serna (2005) (henceforth, LRS) who obtain a transformed GC (TGC) density in order to model conditional higher-order moments based on GARCH dynamics for both skewness and kurtosis. In short, they allow asset return innovations (or standardized errors) to be time-varying TGC (TV-TGC) conditionally distributed. The LRS model, or TV-TGC density, has been used in numerous financial econometric applications due to the increasing interest in modeling conditional higher-order moments. 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).

Unlike the GC density parameters, the TGC density's can no longer be interpreted as the skewness and kurtosis. In this paper, we study the parametric properties of the TGC and obtain the true higher-order moments, which result to be non-linear functions of the GC parameters. Second, we also derive conditions for unimodality, closed-form formulae for (i) the cumulative distribution function (cdf), and (ii) asymmetric risk measures such as expected shortfall (ES) and lower partial moments (LPMs). Third, we illustrate the practical use of this pdf through an application to model asset returns. For that purpose, we implement the threshold GARCH (TGARCH) model of Zakořan (1994) for the conditional volatility together with either constant or TV skewness and kurtosis. We also obtain the closed-form expression for some truncated moments of TGC in order to compute the unconditional variance of the error term under the TGARCH-TGC model.

Fourth, we test the performance of the model through backtesting VaR and ES and the entire density for several types of assets including, stock indexes, exchange rates, commodities and cryptocurrencies. For

comparison purposes, we consider the normal distribution as the benchmark as well as two densities nested in the GC pdf. Namely, the symmetric-GC pdf of Zoia, Biffi and Nicolussi (2018) and the skewed GC (with fixed kurtosis) which we refer to as GCK and GCS, respectively. Unlike the TGC, the GCK and GCS densities do model directly skewness and kurtosis. Density forecasting is evaluated through p-value discrepancy plots along with proper scoring rules (Amisano and Giacomini, 2007). VaR and ES forecasting accuracy are tested through the Du and Escanciano (2017) tests. Our results show that the TGC density performs rather well in forecasting the whole density, VaR and ES, in particular for smaller quantiles (coverage levels of 1% and 2.5% for VaR and ES, respectively). We do not find any differences in performance with regard the type of asset.

Fifth, we also estimate TV higher-order moments according to the TV-TGC conditional distribution where the dynamics for the implied TGC parameters are now driven by the JR (2003) specification with asymmetric responses of conditional skewness and kurtosis to positive and negative shocks.

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 asset returns. Section 5 provides an empirical application to return series. In Section 6 we extend the TGC to allow for time-varying conditional higher-order moments. Section 7 provides a summary of the conclusions. Appendix 1 includes some properties of Hermite polynomials used throughout the paper. Finally, all proofs are included in Appendix 2.

2 The GC distribution

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

$$g(x, \boldsymbol{\theta}) = \phi(x) \psi(x, \boldsymbol{\theta}), \quad (1)$$

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

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

such that $H_k(\cdot)$ denote the (normalized) Hermite polynomials in (54) in Appendix 1. The cdf, i.e. $G(x, \boldsymbol{\theta}) = \int_{-\infty}^x g(u, \boldsymbol{\theta}) du$, is given by

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

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 (henceforth, 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 notations: $\theta_1 = s$ and $\theta_2 = ek$. Since $g(\cdot)$ can take negative values for certain values of (s, ek) , JR

(2001) obtain numerically a restricted space Γ for possible values of (s, ek) where the polynomial function of degree four in (2) becomes non-negative for every x , i.e. $\psi(x, \theta) \geq 0$. As a result, the points in Γ verify that $0 \leq ek \leq 4$, $|s| \leq 1.0493$ and the range of s in Γ depends on the level of ek . The maximum size for skewness is reached for $ek = 2.4508$. From now on, the GC pdf refers to $g(x, \theta)$ subject to $\theta \in \Gamma$. The envelope of Γ is exhibited, in continuous-line, in the left panel of Figure 1 (with ek and s in the x-axis and y-axis, respectively)

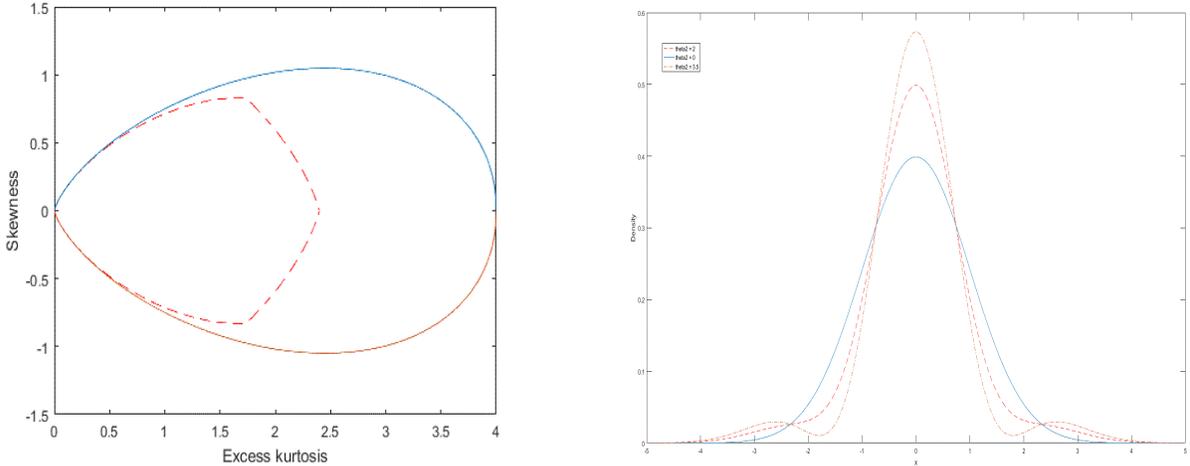


Figure 1: The left panel exhibits the GC positivity frontier (continuous line) and its unimodality frontier (dash line). The right panel plots the GC density for $\theta_1=0$ (symmetric distribution) and different values of excess kurtosis, θ_2 .

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 the 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 Γ which ensures the positivity to the pdf in (1), Gallant and Tauchen (1989) suggested to square the polynomial component $\psi(\cdot, \theta)$, defined in (2), in the pdf $g(\cdot)$. As a result, we can obtain a new pdf $q(\cdot)$ that is called the transformed GC (TGC) density given by

$$q(x, \theta) = \lambda \phi(x) \psi^2(x, \theta), \quad (4)$$

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 contrary to 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 express (4) as

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

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

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

where $\boldsymbol{\nu} = (\nu_0, \nu_1, \dots, \nu_n)' \in \mathbb{R}^{n+1}$ and let $\nu_0 = 1$ 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 $\nu_1 = \nu_2 = 0$.

Finally, let $Q(x, \boldsymbol{\theta}) = \int_{-\infty}^x q(u, \boldsymbol{\theta}) du$ denote the cdf corresponding to the TGC distribution whose closed-form expression is given in the following proposition.

Proposition 1. *Let $Q(x, \boldsymbol{\theta})$ be the cdf of x with $q(\cdot)$ as pdf defined in (5), then*

$$Q(x, \boldsymbol{\theta}) = \lambda \Phi(x) + 2\lambda \gamma_1 \Gamma_{30}(x) + 2\lambda \gamma_2 \Gamma_{40}(x) + 2\lambda \gamma_1 \gamma_2 \Gamma_{34}(x) + \lambda \gamma_1^2 \Gamma_{33}(x) + \lambda \gamma_2^2 \Gamma_{44}(x), \quad (7)$$

such that $\Gamma_{ij}(x) \equiv E_\phi[H_i(u) H_j(u) \mathcal{I}(u \leq x)]$ where $\mathcal{I}(A) = 1 \Leftrightarrow A$ is verified (otherwise, $\mathcal{I}(A) = 0$).

Proof. See Appendix 2. ■

To shorten, $\Gamma_{ij}(\cdot)$ also denotes $\Gamma_i(x)$ in (7) for $j = 0$. We can rewrite (7) as $Q(x, \boldsymbol{\theta}) = \sum_{k=0}^8 \omega_k B_k(x)$ such that $\omega_k = \omega_k(\boldsymbol{\theta})$ are coefficients depending on $\boldsymbol{\theta}$ and $B_k(x) = \int_{-\infty}^x z^k \phi(z) dz$ is in (57) in Appendix 2.

3.1 Higher-order moments

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

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

Proof. See Appendix 2. ■

Let z denote the standardised rv of x , then $z = a(\boldsymbol{\theta}) + b(\boldsymbol{\theta})x$ where $a = -bE_q[x]$, $b = 1/\sigma_x$ and $\sigma_x^2 = E_q[x^2] - E_q[x]^2$. 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

$$s_z \equiv E_q\left[(a + bx)^3\right] = a^3 + 3a^2bE_q[x] + 3ab^2E_q[x^2] + b^3E_q[x^3], \quad (9)$$

$$k_z \equiv E_q\left[(a + bx)^4\right] = a^4 + 4a^3bE_q[x] + 6a^2b^2E_q[x^2] + 4ab^3E_q[x^3] + b^4E_q[x^4]. \quad (10)$$

Figure 2 exhibits the skewness and excess kurtosis region (shaded area) for the TGC distribution given the above equations (9) and (10). 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 $ek_z > 2.7208$ are not captured under the TGC and so, it is less flexible than the GC for more leptokurtic distributions.

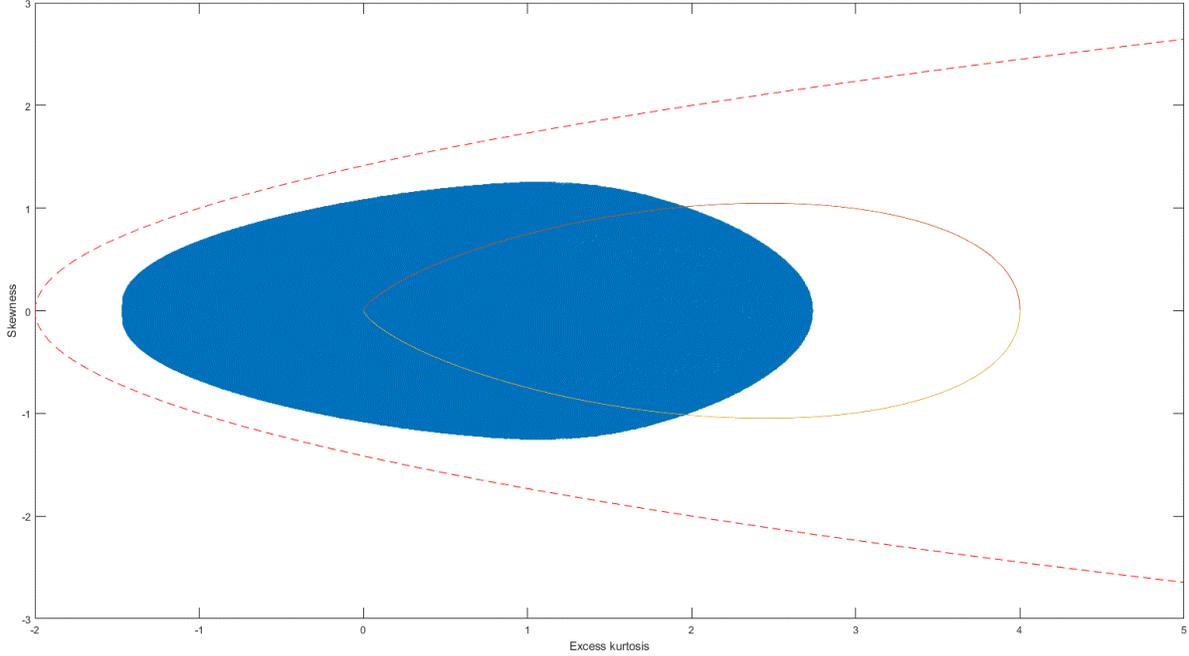


Figure 2: Allowable skewness and excess kurtosis region for TGC pdf (shaded area). The dash-line represents the skewness-excess kurtosis boundary. The GC envelope is represented by the continuous line.

Figure 3 depicts the skewness function (9). 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. $s_z(\theta_1, \bar{\theta}_2)$ where $\bar{\theta}_2 \in \{-3, 0, 3\}$. Note that the graph of $s_z(\theta_1, \bar{\theta}_2)$ behaves like an odd function with respect to θ_1 : $s_z(-\theta_1, \bar{\theta}_2) = -s_z(\theta_1, \bar{\theta}_2)$. The sign of s_z coincides with that of θ_1 for $\bar{\theta}_2 = 0$ as exhibited in the curve $s_z(\theta_1, 0)$. The maximum size of s_z is also obtained for $\bar{\theta}_2 = 0$. The right panel shows the dynamics of s_z as θ_2 varies while θ_1 is fixed, i.e. $s_z(\bar{\theta}_1, \theta_2)$ with selected values for θ_1 just the same as those for θ_2 in the left panel. It is verified that $s_z(0, \theta_2) = 0$ and $s_z(-\bar{\theta}_1, \theta_2) = -s_z(\bar{\theta}_1, \theta_2)$ when $\bar{\theta}_1 \neq 0$ (symmetry respecting the x-axis, denoted as θ_2). We can see that s_z decreases (increases) if θ_2 increases for $\bar{\theta}_1 = 3$ ($\bar{\theta}_1 = -3$).

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 (10). The left and right panels display $ek_z(\theta_1, \bar{\theta}_2)$ and $ek_z(\bar{\theta}_1, \theta_2)$, respectively. Note that in the left panel, the graph $ek_z(\theta_1, \bar{\theta}_2)$ behaves like an even function with respect to θ_1 : $ek_z(-\theta_1, \bar{\theta}_2) = ek_z(\theta_1, \bar{\theta}_2)$. Hence, the sign of θ_1 does not influence the behavior of ek . As a result, we can see in the right panel that $ek_z(\bar{\theta}_1, \theta_2) = ek_z(-\bar{\theta}_1, \theta_2)$ when $\bar{\theta}_1 \neq 0$. Higher values of ek are obtained (in most cases) for $\bar{\theta}_1 = 0$.

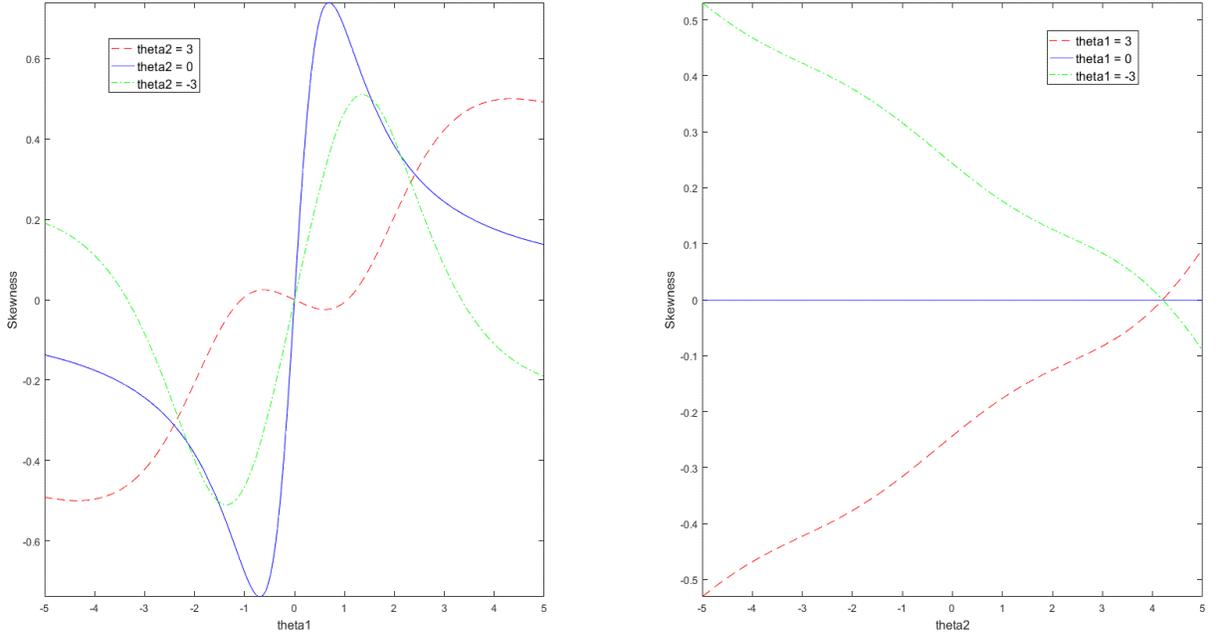


Figure 3: The left panel displays the skewness function $s_z(\theta_1, \bar{\theta}_2)$ for θ_1 given $\bar{\theta}_2 \in \{-3, 0, 3\}$. The right panel shows the range of skewness function $s_z(\bar{\theta}_1, \theta_2)$ for θ_2 given $\bar{\theta}_1 \in \{-3, 0, 3\}$.

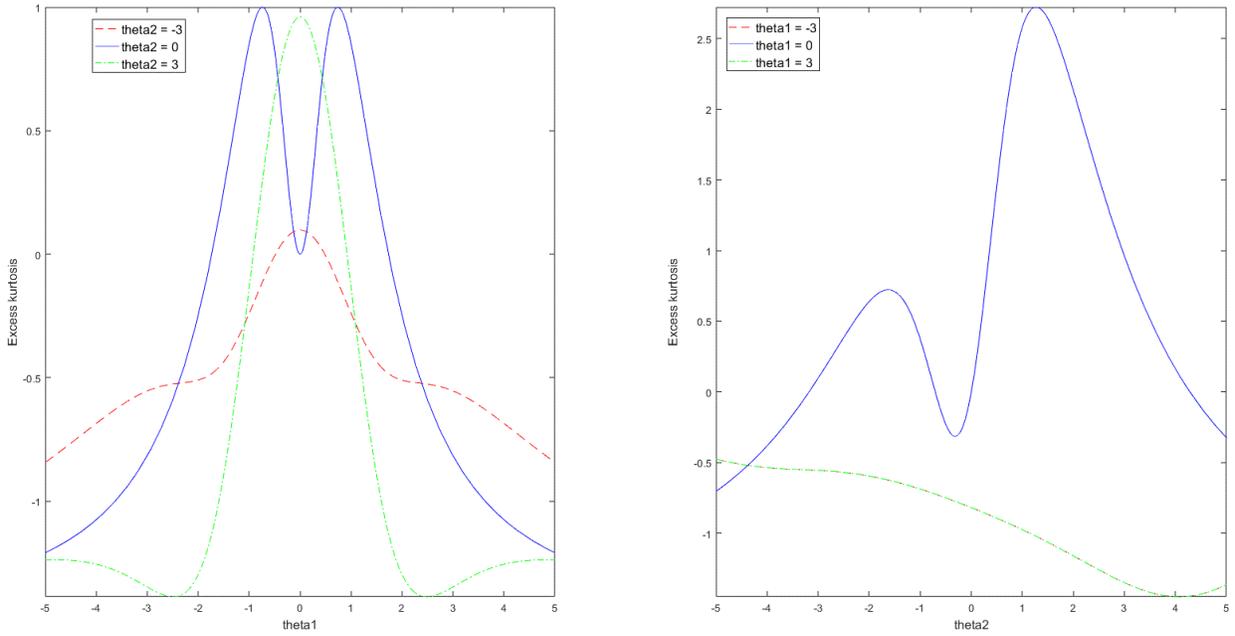


Figure 4: The left and right panels display excess-kurtosis functions $ek_z(\theta_1, \bar{\theta}_2)$ and $ek_z(\bar{\theta}_1, \theta_2)$, for $\bar{\theta}_1, \bar{\theta}_2 \in \{-3, 0, 3\}$, respectively.

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 right 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 left panel in Figure 5 exhibits the values of θ_1 and θ_2 such that the unimodality is verified.

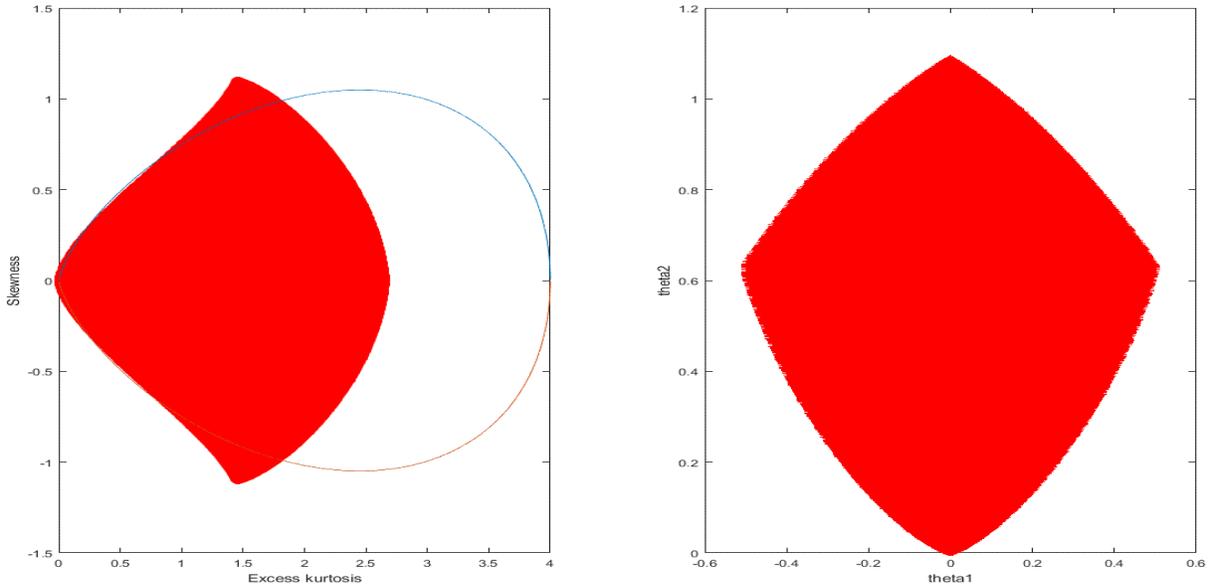


Figure 5: The left panel contains the TGC unimodal region (shaded area) in terms of skewness-excess kurtosis. The right panel exhibits the values of θ_1 and θ_2 such that unimodality holds. The dash line is the usual GC envelope.

Note that (9) and (10) 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 moment series. Specifically, following Amédée-Manesme, Barthélémy and Maillard (2019), we implement a quadratic response surface model, i.e.

$$y = \beta_0 + \beta_1\theta_1 + \beta_2\theta_2 + \beta_3\theta_1^2 + \beta_4\theta_2^2 + \beta_5\theta_1\theta_2 + \varepsilon, \quad (11)$$

where y denotes either s_z or k_z ,¹ 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

¹Because of symmetry in Figure 5 (left panel) respecting the x-axis, we run equation (11) 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

The ES of the random variable x with pdf $q(\cdot)$ in (5) is obtained as $E_q[x|x \leq x_\alpha] = \frac{1}{\alpha} \int_{-\infty}^{x_\alpha} xq(x, \boldsymbol{\theta}) dx$, where $x_\alpha \equiv Q^{-1}(\alpha)$ is the α -quantile (or VaR), i.e. $Q^{-1}(\alpha) = \inf\{x | Q(x, \boldsymbol{\theta}) \geq \alpha\}$.

Proposition 3. *Let $ES_q(\alpha) = E_q[x|x \leq x_\alpha]$ denote the ES with pdf $q(\cdot)$ in (5), then*

$$ES_q(\alpha) = \delta\Gamma_{100}(x_\alpha) + 2\delta\gamma_1\Gamma_{130}(x_\alpha) + 2\delta\gamma_2\Gamma_{140}(x_\alpha) + 2\delta\gamma_1\gamma_2\Gamma_{134}(x_\alpha) + \delta\gamma_1^2\Gamma_{133}(x_\alpha) + \delta\gamma_2^2\Gamma_{144}(x_\alpha), \quad (12)$$

where $\delta \equiv \lambda/\alpha$ and $\Gamma_{ijk}(x) \equiv E_\phi[H_i(u)H_j(u)H_k(u)\mathcal{I}(u \leq x)]$.

Proof. See Appendix 2. ■

To shorten, $\Gamma_{ijk}(\cdot)$ also denotes in (12) both $\Gamma_i(x)$ and $\Gamma_{ij}(\cdot)$ for $j = k = 0$ and $k = 0$, respectively. We can also rewrite (12) as $ES_q(\alpha) = \sum_{k=0}^9 \vartheta_k B_k(x_\alpha)$ where $\vartheta_k = \vartheta_k(\boldsymbol{\theta})$ are coefficients depending on $\boldsymbol{\theta}$ and $B_k(\cdot)$ given in (57) in Appendix 2. Finally, we also obtain the ES under the GC distribution given some results in Proposition 3.

Corollary 1. *Let $ES_g(\alpha) \equiv E_g[x|x \leq x_\alpha]$ be the ES of x with pdf $g(\cdot)$ in (1) and $x_\alpha \equiv G^{-1}(\alpha)$ as the α -quantile with cdf $G(\cdot)$ in (3), then*

$$ES_g(\alpha) = \frac{1}{\alpha}\Gamma_{100}(x_\alpha) + \frac{\gamma_1}{\alpha}\Gamma_{130}(x_\alpha) + \frac{\gamma_2}{\alpha}\Gamma_{140}(x_\alpha), \quad (13)$$

such that $\Gamma_{ijk}(\cdot)$ can be seen in (12).

Proof. Given the results in Proposition 2, the proof is obtained straightforwardly. ■

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

$$LPM_q(\tau, m) = \int_{-\infty}^{\tau} (\tau - x)^m q(x, \boldsymbol{\theta}) dx. \quad (14)$$

The LPM of order 1 is easily obtained as

$$LPM_q(\tau, 1) = \tau Q(\tau, \boldsymbol{\theta}) - Q(\tau, \boldsymbol{\theta}) ES_q(Q(\tau, \boldsymbol{\theta})), \quad (15)$$

such that $ES_q(Q(\tau, \boldsymbol{\theta})) = E_q[x|x \leq \tau]$ is the equation in (12). The LPM of order 2 is given by

$$LPM_q(\tau, 2) = \tau^2 Q(\tau, \boldsymbol{\theta}) - 2\tau Q(\tau, \boldsymbol{\theta}) ES_q(Q(\tau, \boldsymbol{\theta})) + E_q[x^2 \mathcal{I}(x \leq \tau)], \quad (16)$$

where $E_q[x^2 \mathcal{I}(x \leq \tau)]$ is given in the following result.

Corollary 2. *Let $E_q[x^2 \mathcal{I}(x \leq \tau)] = \int_{-\infty}^{\tau} x^2 q(x, \boldsymbol{\theta}) dx$ with $q(\cdot)$ in (5) is obtained as*

$$\begin{aligned} E_q[x^2 \mathcal{I}(x \leq \tau)] &= Q(\tau, \boldsymbol{\theta}) + \sqrt{2}\lambda \Gamma_{200}(\tau) + 2\sqrt{2}\lambda \gamma_1 \Gamma_{230}(\tau) + 2\sqrt{2}\lambda \gamma_2 \Gamma_{240}(\tau) \\ &\quad + 2\sqrt{2}\lambda \gamma_1 \gamma_2 \Gamma_{234}(\tau) + \sqrt{2}\lambda \gamma_1^2 \Gamma_{233}(\tau) + \sqrt{2}\lambda \gamma_2^2 \Gamma_{244}(\tau), \end{aligned} \quad (17)$$

where $Q(\cdot)$ is the cdf in (7) and $\Gamma_{ijk}(x) \equiv E_\phi[H_i(u)H_j(u)H_k(u)\mathcal{I}(u \leq x)]$.

Proof. See Appendix 2. ■

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 Níguez and Perote (2012) with pdf defined as

$$\tilde{q}(x, \boldsymbol{\nu}) = \lambda \phi(x) [1 + \nu_1^2 H_3^2(x) + \nu_2^2 H_4^2(x)], \quad (18)$$

where $\boldsymbol{\nu} = (\nu_1, \nu_2)' \in \mathbb{R}^2$ and $1/\lambda = 1 + \nu_1^2 + \nu_2^2$. The non-central moments are easily obtained in the following result:

Corollary 3. *The first four noncentral moments of x with (18) as pdf are given by*

$$\begin{aligned} E_{\tilde{q}}[x] &= 0, & E_{\tilde{q}}[x^3] &= 0, \\ E_{\tilde{q}}[x^2] &= 6\lambda\nu_1^2 + 8\lambda\nu_2^2 + 1, & E_{\tilde{q}}[x^4] &= 72\lambda\nu_1^2 + 120\lambda\nu_2^2 + 3. \end{aligned} \quad (19)$$

Proof. It is obtained straightforwardly given some results of Proposition 1 in Appendix 2. ■

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

$$k_z = \frac{3(24\lambda\nu_1^2 + 40\lambda\nu_2^2 + 1)}{(6\lambda\nu_1^2 + 8\lambda\nu_2^2 + 1)^2}. \quad (20)$$

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:

$$q_n(x, \nu_n) = \lambda \phi(x) [1 + \nu_n^2 H_n^2(x)], \quad (21)$$

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 4. *Let z be the standardised rv of x with (21) as pdf, i.e. $z = bx$ where $b = 1/\sqrt{E_{q_n}[x^2]}$, then the kurtosis of z is obtained as*

$$k_z(n) = \frac{2\sqrt{6}\lambda\nu_n^2 (A_{4nn} + \sqrt{3}A_{2nn}) + 3}{(\sqrt{2}\lambda\nu_n^2 A_{2nn} + 1)^2}, \quad (22)$$

where $A_{jnn} \equiv E_\phi [H_j(x) H_n^2(x)]$.

Proof. It is obtained straightforwardly given some results of Proposition 1 in Appendix 2. ■

It can be shown that the excess kurtosis obtained from (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 $\mu_t = E[r_t | I_{t-1}]$ and $\sigma_t^2 = E[(r_t - \mu_t)^2 | I_{t-1}]$ 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. $z_t \sim D_t(0, 1)$. Note that D_t nests the simple case of constant parameters across time of the distribution of z_t , i.e. $z_t \sim iid D(0, 1)$. Respecting the conditional variance, we assume the popular TGARCH structure 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

$$r_t = \mu_t + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim D_t(0, 1), \quad (23)$$

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

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

Proposition 4. Let $z_t \sim \text{iid } D(0, 1)$ in (23), then the unconditional variance of ε_t with conditional one driven by the process (24) is

$$\sigma_\varepsilon^2 \equiv E(\sigma_t^2) = \frac{\alpha_0^2(1 + \varpi_1)}{(1 - \varpi_1)(1 - \varpi_2)}, \quad (25)$$

where $\varpi_k = E(c_t^k)$ with $c_t = \beta + \alpha_1^+ z_t^+ - \alpha_1^- z_t^-$ and $\sigma_\varepsilon^2 < +\infty$ for $\varpi_k < 1$, $k = 1, 2$. The expressions for ϖ_k are given by

$$\varpi_1 = \beta - (\alpha_1^- + \alpha_1^+) E(z_t^-), \quad (26)$$

and

$$\varpi_2 = \beta^2 + (\alpha_1^+)^2 + [(\alpha_1^-)^2 - (\alpha_1^+)^2] E[(z_t^-)^2] - 2\beta(\alpha_1^- + \alpha_1^+) E(z_t^-). \quad (27)$$

Proof. See Appendix 2. ■

The second-order stationarity condition of (24) is $\varpi_2 < 1$. The following results show the formulas of (26) and (27) under two particular distributions $D(0, 1)$ for z_t in Proposition 4. Specifically, the first case related to the standard Normal distribution was already obtained in Francq and Zakoian (2010), and the second under the TGC distribution.

Corollary 5. Let $z_t \sim \text{iid } N(0, 1)$ in (23), then the expressions for ϖ_k in Proposition 4 are obtained as

$$\varpi_1 = \beta + \frac{1}{\sqrt{2\pi}} (\alpha_1^- + \alpha_1^+), \quad (28)$$

and

$$\varpi_2 = \beta^2 + \frac{1}{2} [(\alpha_1^-)^2 + (\alpha_1^+)^2] + \sqrt{\frac{2}{\pi}} \beta (\alpha_1^- + \alpha_1^+). \quad (29)$$

Proof. See Appendix 2. ■

Corollary 6. Let $z_t \sim \text{iid } D(0, 1)$ in (23) where $D(0, 1)$ represents the standardized TGC distribution: $z_t = a + bx_t$ such that $x_t \sim \text{TGC}(\boldsymbol{\theta})$ with pdf in (5). Hence, the expressions of $E[(z_t^-)^k]$ for $k = 1, 2$ in Proposition 4 are

$$E(z_t^-) = aQ(-a/b, \boldsymbol{\theta}) + b\lambda\Psi_1(-a/b, \boldsymbol{\theta}), \quad (30)$$

and

$$E[(z_t^-)^2] = (a^2 + b^2) Q(-a/b, \boldsymbol{\theta}) + 2ab\lambda\Psi_1(-a/b, \boldsymbol{\theta}) + \sqrt{2}b^2\lambda\Psi_2(-a/b, \boldsymbol{\theta}), \quad (31)$$

such that $Q(\cdot)$ is the cdf of TGC in (7), $\Psi_k(\cdot)$ is defined as

$$\Psi_k(x, \boldsymbol{\theta}) = \Gamma_{k00}(x) + 2\gamma_1\Gamma_{k30}(x) + 2\gamma_2\Gamma_{k40}(x) + 2\gamma_1\gamma_2\Gamma_{k34}(x) + \gamma_1^2\Gamma_{k33}(x) + \gamma_2^2\Gamma_{k44}(x), \quad (32)$$

where $\Gamma_{kij}(x) \equiv E_\phi[H_k(u)H_i(u)H_j(u)\mathcal{I}(u \leq x)]$, $\gamma_1 = \theta_1/\sqrt{3!}$, $\gamma_2 = \theta_2/\sqrt{4!}$ and $1/\lambda = 1 + \gamma_1^2 + \gamma_2^2$.

Proof. See Appendix 2. ■

4.1 Conditional log-likelihood

The conditional density of r_t can be expressed in terms of the conditional pdf of x_t in (5) as

$$f(r_t | I_{t-1}) = \frac{1}{b_t \sigma_t} q\left(\frac{z_t - a_t}{b_t} | I_{t-1}\right), \quad (33)$$

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

$$\begin{aligned} l_t = & -\ln \sigma_t(\boldsymbol{\varphi}) - \ln b(\boldsymbol{\theta}_t) - \frac{1}{2} \ln(2\pi) + \ln \lambda(\boldsymbol{\theta}_t) \\ & - \frac{1}{2} \left(\frac{z_t(\boldsymbol{\varphi}) - a(\boldsymbol{\theta}_t)}{b(\boldsymbol{\theta}_t)} \right)^2 + \ln \left[\psi \left(\frac{z_t(\boldsymbol{\varphi}) - a(\boldsymbol{\theta}_t)}{b(\boldsymbol{\theta}_t)} \right) \right]^2, \end{aligned} \quad (34)$$

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 $\boldsymbol{\varphi} = (\mu, \alpha_0, \beta, \alpha_1^+, \alpha_1^-)$, and (ii) $\boldsymbol{\theta}_t = \boldsymbol{\theta}$. Finally, this particular case means that (33) can be replaced by the expression: $f(r_t | I_{t-1}) = \frac{1}{b\sigma_t} q\left(\frac{z_t - a}{b}\right)$.

4.2 Conditional asymmetric risk measures

Let $F(r_t | I_{t-1})$ denote the cdf of r_t with the corresponding pdf in (33),

$$F(r_t | I_{t-1}) = \int_{-\infty}^{r_t} f(u | I_{t-1}) du = Q\left(\frac{r_t - \kappa_{0t}}{\kappa_{1t}} | I_{t-1}\right), \quad (35)$$

where $Q(\cdot | I_{t-1})$ is the conditional cdf of $Q(\cdot)$ in (7) 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 $r_{\alpha,t} \equiv F^{-1}(\alpha | I_{t-1})$. Then,

$$r_{\alpha,t} = \kappa_{0t} + \kappa_{1t} Q_t^{-1}(\alpha), \quad (36)$$

such that $Q_t^{-1}(\alpha) \equiv \inf\{x | Q(x | I_{t-1}) \geq \alpha\}$ 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

$$\begin{aligned} ES_t(\alpha) &= E_{t-1}(r_t | r_t \leq r_{\alpha,t}) \\ &= \kappa_{0t} + \kappa_{1t} E_{t-1}(x_t | x_t \leq x_{\alpha,t}), \end{aligned} \quad (37)$$

where $E_{t-1}(x_t | x_t \leq x_{\alpha,t})$ is the conditional version of (12) and $x_{\alpha,t} = (r_{\alpha,t} - \kappa_{0t}) / \kappa_{1t}$ with $r_{\alpha,t}$ as the VaR in (36). Note that $E_{t-1}(\cdot)$ denotes the shortening of $E(\cdot | I_{t-1})$.

The conditional LPM of order 1 and r^* as the threshold of r_t is given by

$$\begin{aligned} LPM_t(r^*, 1) &= \int_{-\infty}^{r^*} (r^* - r_t) f(r_t | I_{t-1}) dr_t \\ &= (r^* - \kappa_{0t}) Q_t(\tau_t) - \kappa_{1t} Q_t(\tau_t) E_{t-1}(x_t | x_t \leq \tau_t), \end{aligned} \quad (38)$$

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

$$\begin{aligned} LPM_t(r^*, 2) &= \int_{-\infty}^{r^*} (r^* - r_t)^2 f(r_t | I_{t-1}) dr_t \\ &= (\tau_t - \kappa_{0t})^2 Q_t(\tau_t) - 2(\tau_t - \kappa_{0t}) \kappa_{1t} Q_t(\tau_t) E_{t-1}(x_t | x_t \leq \tau_t) \\ &\quad + E_{t-1}[x_t^2 \mathcal{I}(x_t \leq \tau_t)], \end{aligned} \quad (39)$$

such that $E_{t-1}[x_t^2 \mathcal{I}(x_t \leq \tau_t)] = \int_{-\infty}^{\tau_t} x_t^2 q(x_t | I_{t-1}) dx_t$ is obtained in (17).

5 Empirical application

5.1 Dataset and modeling

The data used are daily percentage 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.

Initially, we adopt some density functions with constant parameters across time for the conditional standardized returns in (23), i.e. $z_t \sim iid D(0, 1)$. We consider the following cases: (i) the standard Normal distribution, i.e. $D(0, 1) = N(0, 1)$; (ii) the GC-skewed (GCS hereafter) distribution that is nested in the GC pdf in (1) with skewness parameter s and \bar{ek} as the fixed excess kurtosis parameter, i.e. $D(0, 1) = GC(s, \bar{ek})$; (iii) 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 (iv) $D(0, 1)$ is the standardized TGC distribution denoted as $TGC(0, 1, \theta)$, i.e. $z_t = a + bx_t$ such that $x_t \sim TGC(\theta)$ 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).

For the backtesting procedures below, we take the first $T - N$ observations for the first in-sample window and an out-of-sample (OOS) period of length $N = 1,000$ using a daily rolling constant-sized 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, GCS and GCK density parameters are obtained by ML using the standardized residuals, z_t , from the first stage. According to the first in-sample TGARCH parameter estimates in Table 2, the returns for all assets exhibit volatility clustering and asymmetric response to good and bad news. We find that for the Bitcoin return series the estimates differ significantly from those of the other series reflecting rather different volatility patterns. All series verify the second-order stationarity condition, i.e. the estimate of ϖ_2 in (29) is lower than one. The following results can be extracted from Table 3. First, the TGC model yields slightly higher log-likelihood (LL) values than the GCS except for US-UK, and both models are higher than that of the GCK for all series except for Bitcoin. Second, both TGC and GCK models yield similar levels of excess kurtosis. Note that the excess kurtosis of GCS is fixed to the GCK's estimated value. Third, both TGC and GCS produce significant and similar levels of skewness for all series, except for the Bitcoin where the former yields a small positive skewness whilst the latter provides a non-significant negative one.² Overall, the gap in LL values can be mainly due to differences in the estimated levels of the higher-order moments produced by the models.

5.2 Backtesting the density

5.2.1 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 $\{u_t\}_{t=1}^N \sim iid U(0,1)$ under the correct one-step ahead cdf specification with $u_t = D(r_t | I_{t-1})$ where $D(\cdot | I_{t-1})$ 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

$$\widehat{P}_{p_t}(y_\varrho) = \frac{1}{N} \sum_{t=1}^N \mathcal{I}(u_t \leq y_\varrho), \quad (40)$$

where $\mathcal{I}(u_t \leq y_\varrho)$ is an indicator function and y_ϱ is an arbitrary grid of ϱ points.³ 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 the p-value discrepancy plots for all models and series. A first observation that emerges from the plots is that the three GC models (TGC, GCS and GCK) perform overall better than the Normal. We find that PITs from both TGC and GCS tend to cluster and depart from the GCK in most cases. Note that for the Bitcoin there are no visually noticeable differences between the GC models and they exhibit a great difference respecting the Normal.

²We have also estimated the GC with pdf in (1) and found similar parameter estimates as the ones in Table 3 for the TGC.

³We use the following $\varrho = 215$ points grid: $y_\varrho \in \{0.001, 0.002, \dots, 0.01, 0.015, \dots, 0.99, 0.991, \dots, 0.999\}$, since it highlights the goodness-of-fit in the distribution tails.

Table 1: Summary statistics for daily percent log returns

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
Mean	0.0084	0.0022	-0.0014	-0.0048	0.0271	0.3887
Median	0.0166	0.0162	0.0047	0.0000	0.0000	0.2032
Max	11.6442	11.9653	3.7102	4.4744	16.4137	42.4580
Min	-11.1856	-11.1024	-6.5818	-8.3120	-17.0918	-49.1440
Std. dev.	1.4571	1.5600	0.6837	0.5823	2.3749	5.7352
Skewness	-0.2348	-0.0848	-0.4726	-0.5835	-0.1703	-0.3196
Kurtosis	7.4453	8.3963	8.1029	14.4818	7.2350	14.9675
J-B stat	4362.7	7563.7	5838.7	28952.9	3923.8	17564.6
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.

Table 2: QML estimation results

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
μ	-0.0097	0.0064	0.0063	-0.0024	0.0257	0.3182*
	(0.019)	(0.018)	(0.009)	(0.007)	(0.030)	(0.133)
α_0	0.0404*	0.0284*	0.0095*	0.0044*	0.0159*	0.4560*
	(0.0089)	(0.0056)	(0.0036)	(0.0014)	(0.0068)	(0.1939)
β	0.9052*	0.9244*	0.9437*	0.9575*	0.9553*	0.7024*
	(0.0119)	(0.0092)	(0.0100)	(0.0064)	(0.0100)	(0.0712)
α_1^+	0.0409*	0.0143*	0.0439*	0.0323*	0.0296*	0.3332*
	(0.0097)	(0.0072)	(0.0079)	(0.0066)	(0.0095)	(0.0708)
α_1^-	0.1329*	0.1290*	0.0660*	0.0561*	0.0714*	0.3382*
	(0.0165)	(0.0147)	(0.0128)	(0.0077)	(0.0145)	(0.0861)
ϖ_2	0.9546	0.9686	0.9765	0.9864	0.9926	0.9823
LL	-1.7406	-1.7240	-0.9369	-0.7542	-2.1753	-2.9970

Model parameter specifications: $r_t = \mu + \varepsilon_t$, $\varepsilon_t = \sigma_t z_t$, $z_t \sim iid N(0, 1)$, $\sigma_t = \alpha_0 + \beta\sigma_{t-1} + \alpha_1^+ \varepsilon_{t-1}^+ - \alpha_1^- \varepsilon_{t-1}^-$.

This table presents QML estimates of parameters of the TGARCH model for percent log return series: Nikkei, Eurostoxx50, JAP-US, US-UK, Oil and Bitcoin. The in-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns comprises 4,217 observations from January 15, 1999 to March 16, 2015; and for Bitcoin returns comprises 1,935 observations from July 19, 2010 to November 3, 2015. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*) and (**) indicate significance at 1% and 5% levels, respectively. The second-order stationarity condition of TGARCH under normality must verify $\varpi_2 < 1$ with ϖ_2 given in (29). Finally, LL is the log-likelihood value (constant terms included) of the model.

Table 3: TGC, GCS and GCK models estimation results

	TGC	GCS	GCK	TGC	GCS	GCK
	Nikkei			Eurostoxx50		
θ_1	-0.0553*	-0.1239*		-0.0589*	-0.1481*	
	(0.018)	(0.046)		(0.018)	(0.044)	
θ_2	0.2387*		0.5628*	0.2544*		0.5918*
	(0.036)		(0.093)	(0.036)		(0.094)
s	-0.1446			-0.1559		
ek	0.6326			0.6819		
LL	-1.4105	-1.4115	-1.4124	-1.4109	-1.4114	-1.4127
	JAP-US			US-UK		
θ_1	-0.0349*	-0.0942**		-0.0335	-0.0954*	
	(0.018)	(0.048)		(0.018)	(0.045)	
θ_2	0.4319*		1.0960*	0.2295*		0.5266*
	(0.032)		(0.098)	(0.037)		(0.095)
s	-0.1033			-0.0871		
ek	1.2439			0.5980		
LL	-1.3954	-1.3977	-1.3982	-1.4124	-1.4123	-1.4128
	Oil			Bitcoin		
θ_1	-0.0767*	-0.2025*		0.0021	-0.045	
	(0.018)	(0.048)		(0.022)	(0.580)	
θ_2	0.3741*		0.9165*	0.9186*		2.6972*
	(0.032)		(0.095)	(0.071)		(0.119)
s	-0.2192			0.0062		
ek	1.0645			2.4726		
LL	-1.3963	-1.3995	-1.4015	-1.3039	-1.3027	-1.3028

This table presents ML estimates of parameters of the TGC, GCS and GCK models for percent log return series: Nikkei, Eurostoxx50, JAP-US, US-UK, Oil and Bitcoin. The in-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns comprises 4,217 observations from January 15, 1999 to March 16, 2015; and for Bitcoin returns comprises 1,935 observations from July 19, 2010 to November 3, 2015. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*) and (**) indicate significance at 1% and 5% levels, respectively. LL gives log likelihood values (constant terms included); and s and ek denotes in-sample skewness and excess kurtosis estimates from equations (9) and (10), respectively.

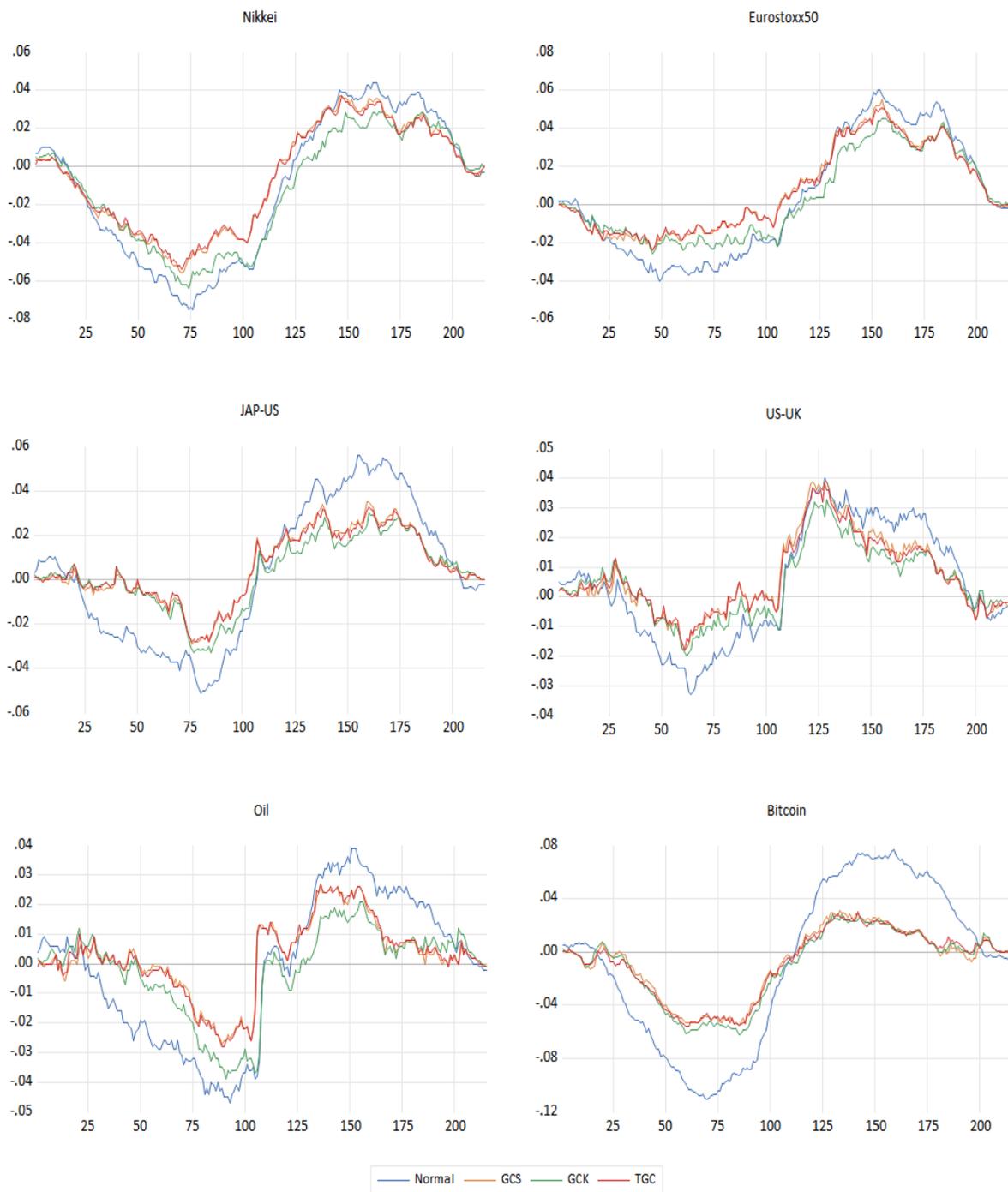


Figure 6: P-value discrepancy plots for PIT series obtained from TGC, GCS, GCK and Normal pdfs. Return series: Nikkei, Eurostoxx50, JAP-US, UK-US, Oil and Bitcoin. The out-of-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns covers from March 17, 2015 to January 14, 2019; and for Bitcoin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

5.2.2 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 $\Upsilon(\tilde{f}_t, r_t)$ whose arguments are the density forecast $\tilde{f}_t = f_{t-1}(r_t)$ and the future realization of the return, r_t . In this paper, we use the weighted logarithmic scoring rule:

$$\Upsilon_j(\tilde{f}_t, z_t) = \omega_j(z_t) \ln \tilde{f}_t, \quad (41)$$

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:

$$\bar{\Upsilon}_j(\tilde{f}_t, z_t) = N^{-1} \sum_{t=1}^N \Upsilon_{jt}(\tilde{f}_t, z_t). \quad (42)$$

So, we prefer model f if $\bar{\Upsilon}_j(\tilde{f}_t, z_t) > \bar{\Upsilon}_j(\tilde{g}_t, z_t)$, and prefer model g otherwise. The null hypothesis $H_0 : E \left[\bar{\Upsilon}_j(\tilde{f}_t, z_t) - \bar{\Upsilon}_j(\tilde{g}_t, z_t) \right] = 0$ is tested in Amisano and Giacomini (2007).

Table 4: 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
GCS	left	-0.6431	-0.6970	-0.7143	-0.7785	-0.7143	-0.6431
	center	-0.2943	-0.3317	-0.3192	-0.3233	-0.3192	-0.2943
	right	-0.6437	-0.6442	-0.6595	-0.6977	-0.6595	-0.6437
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 (42) for one-step-ahead density forecast from TGC, GCS, GCK and Normal (N) 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.

Table 4 presents the results of the weighted average scores. A first observation that stands out is that all TGC's weighted scores are higher than those of the rest of the models for all series. The Normal model provides systematically the lowest scores. Both GCK and GCS lead to similar performance and in-between the TGC and the Normal. We carried out Amisano and Giacomini (2007) tests, which show that differences in models' scores are significant for almost all cases.⁴

5.3 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

$$VaR_t(\alpha) = \kappa_{0,t} + \kappa_{1,t}Q^{-1}(\alpha), \quad (43)$$

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, which incorporates the exception magnitude, provides useful information to discriminate among similar models according to the unconditional coverage test and is given by

$$QL_t(\alpha) = (r_t - y_t)^2 \times h_t(\alpha), \quad (44)$$

where $y_t \in \{VaR_t(\alpha), ES_t(\alpha)\}$, see Lopez (1999) and Angelidis and Degiannakis (2007). The sample average of (44) for the OOS period of N days is

$$AQL(\alpha) = N^{-1} \sum_{t=1}^N QL_t(\alpha). \quad (45)$$

5.3.1 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.

$$U_{VaR}(\alpha) = \frac{\sqrt{N}(\bar{h}(\alpha) - \alpha)}{\sqrt{\alpha(1-\alpha)}} \stackrel{a}{\approx} N(0, 1), \quad (46)$$

where $\bar{h}(\alpha)$ is the sample average of $\{\hat{h}_t(\alpha)\}_{t=1}^N$ such that $\hat{h}_t(\alpha) = \mathcal{I}(\hat{u}_t \leq \alpha)$ with \hat{u}_t as the estimation of $u_t = F(r_t | I_{t-1})$ in (35). 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:

$$C_{VaR}(m) = N \sum_{i=1}^m \hat{\rho}_i^2 \stackrel{a}{\approx} \chi_m^2, \quad (47)$$

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 $\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0}$ where

$$\hat{\gamma}_j = \frac{1}{N-j} \sum_{t=1+j}^N (\hat{h}_t(\alpha) - \alpha) (\hat{h}_{t-j}(\alpha) - \alpha). \quad (48)$$

⁴These test results are not included to save space but are available from the authors upon request.

5.3.2 Backtesting ES

The unconditional and conditional backtests are the analogues for ES of the above unconditional and conditional VaR backtests. 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

$$\begin{aligned}\mathcal{H}_t(\alpha) &= \int_0^\alpha h_t(u) du \\ &= \frac{1}{\alpha} (\alpha - u_t) \mathcal{I}(u_t \leq \alpha).\end{aligned}\quad (49)$$

Note that (49) measures the distance of the returns from the corresponding α -quantile in (43) during the violations. It is shown that $\{\mathcal{H}_t(\alpha) - \frac{\alpha}{2}\}_{t=1}^\infty$ follows the MDS property. Thus, the null hypothesis for the unconditional backtest is $H_{0,U} : E[\mathcal{H}_t(\alpha)] = \frac{\alpha}{2}$ and the related test statistics is given by

$$U_{ES} = \frac{\sqrt{N}(\overline{\mathcal{H}}(\alpha) - \frac{\alpha}{2})}{\sqrt{\alpha(\frac{1}{3} - \frac{\alpha}{4})}} \stackrel{a}{\sim} N(0, 1), \quad (50)$$

where $\overline{\mathcal{H}}(\alpha)$ is the mean of $\{\widehat{\mathcal{H}}_t(\alpha)\}_{t=1}^N$ 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 $H_{0,C} : E[\mathcal{H}_t(\alpha) | I_{t-1}] = \frac{\alpha}{2}$ with corresponding test statistics the Box-Pierce one given by

$$C_{ES}(m) = N \sum_{i=1}^m \hat{\rho}_j^2 \stackrel{a}{\sim} \chi_m^2, \quad (51)$$

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

$$\hat{\gamma}_j = \frac{1}{N-j} \sum_{t=1+j}^N \left(\widehat{\mathcal{H}}_t(\alpha) - \frac{\alpha}{2} \right) \left(\widehat{\mathcal{H}}_{t-j}(\alpha) - \frac{\alpha}{2} \right). \quad (52)$$

5.3.3 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 5 shows the results of the descriptive analysis of violations. Firstly, we find that for the low coverage levels suggested by the Basel Committee (VaR(0.01), ES(0.025)) all three GC models perform much better than the Normal with the TGC and the GCS providing the best performance for all series. For 5% VaR and 10% ES differences are smaller between GC and Normal models and the latter performs better for Bitcoin. These results are reinforced by the magnitude of exceptions for VaR and ES measured through the $AQL(\alpha)$ statistic presented in Table 6. We find that either TGC or GCS yield AQL lower values for 1% and 2.5% coverages, and for 5% (but not 10%) level the Normal distribution performs better than the GCK. Table 7 shows the results of the p-values of VaR and ES backtesting. The null is accepted for most return series at the low coverage levels pointing out we cannot find empirical evidence mainly against all three GC densities.

Table 5: 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	37.89
GCS	14	15.34	43	42.78	7	8.48	40	37.17
GCK	18	17.40	46	45.75	7	9.68	41	40.17
N	18	19.40	46	45.26	12	11.49	41	38.90
	JAP-US				US-UK			
TGC	12	13.45	53	50.55	10	14.70	54	54.14
GCS	12	13.76	51	49.43	10	14.69	50	52.93
GCK	12	14.82	54	51.34	11	15.85	56	55.62
N	19	20.60	52	49.48	18	19.56	54	53.81
	Oil				Bitcoin			
TGC	13	12.01	53	52.45	6	7.60	51	45.43
GCS	13	11.91	51	51.41	6	6.82	51	46.92
GCK	13	14.05	56	55.33	6	7.22	54	48.19
N	16	18.36	54	53.45	16	18.09	47	44.48

This table shows the violations for VaR and the cumulative violations given in (49) for ES under each pdf. 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 6: Average quadratic loss of VaR and ES

	VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)	VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)
	Nikkei				Eurostoxx50			
TGC	0.0186	0.0182	0.0725	0.0581	0.0554	0.0549	0.0925	0.0831
GCS	0.0185	0.0183	0.0713	0.0577	0.0553	0.0551	0.0917	0.0829
GCK	0.0229	0.0223	0.0772	0.0634	0.0585	0.0581	0.0951	0.0860
N	0.0279	0.0274	0.0761	0.0653	0.0625	0.0621	0.0943	0.0873
	JAP-US				US-UK			
TGC	0.0055	0.0055	0.0214	0.0166	0.0430	0.0427	0.0601	0.0558
GCS	0.0058	0.0058	0.0208	0.0165	0.0429	0.0427	0.0596	0.0556
GCK	0.0062	0.0062	0.0218	0.0173	0.0436	0.0434	0.0605	0.0564
N	0.0089	0.0087	0.0210	0.0183	0.0457	0.0455	0.0601	0.0570
	Oil				Bitcoin			
GCT	0.0134	0.0133	0.1288	0.0881	0.1251	0.1319	0.7595	0.5224
GCS	0.0135	0.0137	0.1235	0.0868	0.1181	0.1220	0.7988	0.5194
GCK	0.0189	0.0187	0.1405	0.0998	0.1261	0.1311	0.8663	0.5503
N	0.0326	0.0318	0.1345	0.1075	0.2872	0.2832	0.7235	0.6144

This table exhibits the results of the AQL function given in (45) for VaR and ES from alternative 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.

Table 7: P-values for backtesting VaR and ES

		VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)	VaR(1%)	ES(2.5%)	VaR(5%)	ES(10%)
		Nikkei				Eurostoxx50			
TGC	<i>U</i>	0.824	0.204	0.212	0.685	0.035	0.340	0.036	0.043
	<i>C</i>	0.920	0.973	0.533	0.770	0.168	0.999	0.210	0.585
GCS	<i>U</i>	0.809	0.204	0.175	0.685	0.025	0.340	0.023	0.043
	<i>C</i>	0.922	0.973	0.540	0.770	0.193	0.999	0.180	0.585
GCK	<i>U</i>	0.583	0.026	0.469	1.000	0.119	0.340	0.023	0.043
	<i>C</i>	0.935	0.716	0.558	0.782	0.193	0.999	0.180	0.585
N	<i>U</i>	0.319	0.011	0.393	0.685	0.217	0.525	0.046	0.156
	<i>C</i>	0.949	0.759	0.549	0.773	0.078	0.322	0.158	0.753
		JAP-US				US-UK			
TGC	<i>U</i>	0.837	0.525	0.921	0.685	0.506	1.000	0.456	0.685
	<i>C</i>	0.708	0.983	0.200	0.672	0.493	0.122	0.766	0.443
GCS	<i>U</i>	0.997	0.525	0.918	0.685	0.594	1.000	0.598	0.685
	<i>C</i>	0.776	0.525	0.918	0.672	0.486	0.122	0.760	0.443
GCK	<i>U</i>	0.653	0.525	0.770	0.685	0.310	0.751	0.312	0.311
	<i>C</i>	0.670	0.983	0.199	0.672	0.524	0.217	0.777	0.436
N	<i>U</i>	0.180	0.004	0.925	0.311	0.127	0.011	0.492	0.105
	<i>C</i>	0.717	0.952	0.210	0.866	0.560	0.529	0.786	0.447
		Oil				Bitcoin			
TGC	<i>U</i>	0.937	0.340	0.680	0.418	0.178	0.204	0.410	0.068
	<i>C</i>	0.435	0.423	0.006	0.723	0.854	1.000	0.137	0.665
GCS	<i>U</i>	0.738	0.340	0.820	0.418	0.324	0.204	0.745	0.026
	<i>C</i>	0.473	0.423	0.007	0.723	0.760	1.000	0.138	0.474
GCK	<i>U</i>	0.475	0.340	0.334	0.839	0.359	0.204	0.698	0.043
	<i>C</i>	0.345	0.423	0.006	0.821	0.751	1.000	0.126	0.585
N	<i>U</i>	0.157	0.057	0.552	0.311	0.348	0.057	0.320	0.311
	<i>C</i>	0.375	0.662	0.007	0.647	0.732	0.664	0.087	0.436

This table reports the p-values for (i) the VaR backtesting tests in (46) and (47), and (ii) the ES backtesting tests in (50) and (51). The unconditional and conditional backtests are denoted, respectively, as *U* and *C*. 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 Time-varying conditional higher-order moments

The conditional skewness and kurtosis of r_t are defined, respectively, as $s_{r,t} = E(\varepsilon_t^3 | I_{t-1}) / \sigma_t^3$ and $k_{r,t} = E(\varepsilon_t^4 | I_{t-1}) / \sigma_t^4$. If we let the pdf of x_t in (5) 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 higher-order moments where $s_{z,t}$ and $k_{z,t}$ are obtained by plugging $\theta_{i,t}$ into equations (9) and (10). The dynamics equation for $\theta_{i,t}$ is given by

$$\theta_{i,t} = \varphi_{0i} + \varphi_{1i}\theta_{i,t-1} + \varphi_{2i}^+ z_{t-1}^+ + \varphi_{2i}^- z_{t-1}^-. \quad (53)$$

This specification was initially proposed by JR (2003) under the conditional generalized Student's t distribution of Hansen (1994) for the innovations. It has later been implemented, among others, by Feunou, Jahan-Parvar and Tédongap (2016) who consider alternative conditional densities and introduce the skewed generalized error (SGE) distribution; Anatolyev and Petukhov (2016) use the SGE density of Feunou et al. (2016) assuming time variation only for the parameter more related to skewness; Lalancette and Simonato (2017) use the conditional Johnson S_u distribution, and León and Níguez (2018) exploit the conditional SNP distribution in LMS (2009).

Table 8 exhibits the ML parameter estimates for $\theta_{i,t}$ in (53). Several features can be observed. First, there is evidence that the TV-TGC model does capture skewness and kurtosis clustering under the JR asymmetric model since φ_{1i} and φ_{2i} are significant in most returns series for both $\theta_{1,t}$ and $\theta_{2,t}$. Second, we can find some evidence of asymmetric response to positive and negative shocks only for one TV parameter as exhibited in Nikkei and Bitcoin for $\theta_{2,t}$ and $\theta_{1,t}$, respectively. Third, it is also shown that only φ_{2i}^+ is significant in most returns series.

7 Conclusions

In this article we develop a new pdf based on the GC expansion in JR (2001). Our approach consists of a transformation of the GC pdf through the method of Gallant and Nychka (1987) which ensures the pdf's positivity within the whole parametric space. We provide an analysis of the TGC's statistical properties including conditions for unimodality, closed-form expressions of cdf, moments, quantiles, ES and lower partial moments. As an extension, we present a TGC specification with time-varying skewness and kurtosis which gathers clustering and asymmetric response to positive and negative shocks.

The relative performance of our model is tested through an OOS application to forecast the density, VaR and ES of stock indexes, exchange rate, oil and cryptocurrency returns. To do so, and in order to isolate the effect of skewness, we consider the symmetric-GC of Zoia et al. (2018) and its asymmetric extension, which we name GCS and GCK, respectively. We use the Normal distribution as a benchmark for our comparative analysis. Density forecasting is evaluated through p-value discrepancy plots and proper scoring rules in Amisano and Giacomini (2007). VaR and ES forecast accuracy is evaluated through the backtesting methods of Escanciano and Olmo (2010) and Du and Escanciano (2017), respectively.

Our backtesting analysis to forecast the entire distribution shows that GC models (TGC, GCS, GCK) provide a closer fit than the Normal for all returns series. Both TGC and GCS perform similarly and the GCK's fit is between the former and the Normal. In particular, for the distribution center the TGC and GCS fit is very similar and visually different to that of both the GCK and the Normal.

Regarding the backtesting for the tail distribution, the TGC clearly overperforms the GCK and Normal models for lower coverage levels, i.e. 1% for VaR and 2.5% for ES. In comparison to the GCS, the TGC's performance differ slightly and in a minority of cases the GCS's is better.

The GC is an example of a polynomial expansion (PE) density with the standard normal as the parent distribution. Following the methodology of Bagnato et al. (2015), an interesting avenue for future research would be obtaining the PE density according to the standardized Student-t distribution and analyze its parametric properties.

Table 8: TV-TGC (JR asymmetric) model estimation results

	Nikkei	Eurostoxx50	JAP-US	US-UK	Oil	Bitcoin
φ_{01}	-0.0944* (0.034)	-0.0204 (0.029)	-0.0302 (0.026)	-0.0620 (0.036)	-0.1550* (0.034)	0.0402* (0.010)
φ_{11}	-0.4714* (0.213)	0.2349 (0.268)	0.5032* (0.154)	-0.5949* (0.165)	-0.8118* (0.124)	0.8594* (0.069)
φ_{21}^+	0.0608* (0.028)	-0.0091 (0.034)	0.0698** (0.032)	0.0324 (0.030)	0.0455** (0.022)	-0.0461* (0.016)
φ_{21}^-	0.0227 (0.022)	0.0346 (0.032)	0.0324 (0.027)	0.0160 (0.029)	-0.0126 (0.017)	0.0801* (0.024)
φ_{02}	0.3059* (0.099)	-0.0128 (0.047)	0.1612 (0.109)	0.4633* (0.074)	0.0986 (0.077)	1.4579* (0.142)
φ_{12}	-0.6026* (0.233)	0.5882* (0.181)	0.3422 (0.298)	-0.9678* (0.012)	0.5214* (0.196)	-0.4535* (0.135)
φ_{22}^+	0.1237** (0.061)	0.3043* (0.078)	0.1008 (0.061)	-0.0343 (0.019)	0.1899* (0.059)	-0.3328* (0.085)
φ_{22}^-	-0.1022* (0.041)	-0.0546 (0.051)	-0.2021* (0.050)	0.0154 (0.019)	-0.0275 (0.042)	0.0134 (0.075)
LL	-1.4085	-1.4079	-1.3922	-1.4116	-1.3937	-1.2950
$\min(sk)$	-0.6286	-0.5824	-0.5891	-0.4883	-0.6763	-1.1553
$\max(sk)$	0.5112	-0.0620	0.9788	0.2702	0.3724	0.5144
$\text{mean}(sk)$	-0.1363	-0.1353	-0.0967	-0.0915	-0.2047	0.0000
$\min(ku)$	2.8469	3.0016	3.6495	2.9055	3.5720	2.6981
$\max(ku)$	4.9967	5.7026	5.5925	4.5337	5.6900	5.7194
$\text{mean}(ku)$	3.6696	3.8375	4.1819	3.5959	4.0672	5.3741

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 φ_{ki} in the previous parameter equations of $\theta_{i,t}$ for percent log return series: Nikkei, Eurostoxx50, JAP-US, US-UK, Oil and Bitcoin. The "JR asymmetric" model for $\theta_{i,t}$ is borrowed from JR (2003). The in-sample period for Nikkei, Eurostoxx50, JAP-US, US-UK and Oil returns comprises 4,217 observations from January 15, 1999 to March 16, 2015; and for Bitcoin returns comprises 1,935 observations from July 19, 2010 to November 3, 2015. Heteroscedasticity-consistent standard errors are provided in parentheses below the parameter estimates. (*) and (**) indicate significance at 1% and 5% levels, respectively. sk and ku denote in-sample skewness and kurtosis. $\min(x)$, $\max(x)$ and $\text{mean}(x)$ denote minimum, maximum and mean values of series x . LL gives log likelihood values (constant terms included) of the returns model: TGARCH model with constant mean and the above TV-TGC distribution for the standardized returns.

Appendix 1 (Hermite polynomials)

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

$$H_k(x) = \frac{xH_{k-1}(x) - \sqrt{k-1}H_{k-2}(x)}{\sqrt{k}}, \quad (54)$$

with initial conditions $H_0(x) = 1$ and $H_1(x) = x$. The set $\{H_k(x)\}_{k \in \mathbb{N}}$ 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):

$$H_k(x) = \sqrt{k!} \sum_{n=0}^{\lfloor k/2 \rfloor} \frac{(-1)^n}{n!(k-2n)!2^n} x^{k-2n}, \quad (55)$$

where $\lfloor \cdot \rfloor$ rounds its argument to the nearest integer toward zero. Given (55), we can now write x^k in terms of $H_k(x)$ as follows:

$$x^k = \sqrt{k!}H_k(x) - \sqrt{k!} \sum_{n=1}^{\lfloor k/2 \rfloor} \frac{(-1)^n}{n!(k-2n)!2^n} x^{k-2n}. \quad (56)$$

If we substitute recursively the powers of x on the right side of equation (56) 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)

Consider $m_k(x) \equiv E_\phi[z|z \leq x]$, then a recursive formula for the truncated normal moments (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

$$B_k(x) = (k-1)B_{k-2}(x) - x^{k-1}\phi(x), \quad k \geq 2 \quad (57)$$

where $B_0(x) = \Phi(x)$ and $B_1(x) = -\phi(x)$. The recursion formula in (57) is also obtained, in a very slightly different way, in Skoulakis (2019) and it will be useful for all our proofs.

Proof of Proposition 1 (cdf). Consider $\xi_{ijk} \equiv 1/\sqrt{i!j!k!}$, the equation in (55) and $B_k(x)$ in (57), then the expressions $\Gamma_{ij}(\cdot)$ in (7) are easily computed as

$$\begin{aligned} \Gamma_{30}(x) &= \xi_{300} [B_3(x) - 3B_1(x)], \\ \Gamma_{40}(x) &= \xi_{400} [B_4(x) - 6B_2(x) + 3B_0(x)], \\ \Gamma_{33}(x) &= \xi_{330} [B_6(x) - 6B_4(x) + 9B_2(x)], \\ \Gamma_{34}(x) &= \xi_{340} [B_7(x) - 9B_5(x) + 21B_3(x) - 9B_1(x)], \\ \Gamma_{44}(x) &= \xi_{440} [B_8(x) - 12B_6(x) + 42B_4(x) - 36B_2(x) + 9B_0(x)]. \end{aligned} \quad (58)$$

Proof of Proposition 2 (moments). According to (56), 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 $E_q[x^k] = \int_{-\infty}^{+\infty} x^k q(x) dx$ as follows:

$$\begin{aligned}
E_q[x] &= 2\lambda\gamma_1\gamma_2 A_{134}, \\
E_q[x^2] &= \sqrt{2}\lambda\gamma_1^2 A_{233} + \sqrt{2}\lambda\gamma_2^2 A_{244} + 1, \\
E_q[x^3] &= 2\sqrt{3!}\lambda\gamma_1 + 2\sqrt{3!}\lambda\gamma_1\gamma_2 A_{334} + 6\lambda\gamma_1\gamma_2 A_{134}, \\
E_q[x^4] &= 2\sqrt{4!}\lambda\gamma_2 + \sqrt{4!}\lambda\gamma_1^2 A_{334} + \sqrt{4!}\lambda\gamma_2^2 A_{444} + 6\sqrt{2}\lambda\gamma_1^2 A_{233} + 6\sqrt{2}\lambda\gamma_2^2 A_{244} + 3,
\end{aligned} \tag{59}$$

where $A_{ijk} \equiv E_\phi[H_{ijk}(x)] = \int_{-\infty}^{+\infty} H_{ijk}(x) \phi(x) dx$ with $H_{ijk}(x) \equiv H_i(x)H_j(x)H_k(x)$ verifying that $A_{ijk} = 0$ if $i + j + k$ is an odd number and

$$\begin{aligned}
A_{233} &= \xi_{233} [n_8 - 7n_6 + 15n_4 - 9n_2], \\
A_{134} &= \xi_{134} [n_8 - 9n_6 + 21n_4 - 9n_2], \\
A_{244} &= \xi_{244} [n_{10} - 13n_8 + 54n_6 - 78n_4 + 45n_2 - 9], \\
A_{334} &= \xi_{334} [n_{10} - 12n_8 + 48n_6 - 72n_4 + 27n_2], \\
A_{444} &= \xi_{444} [n_{12} - 18n_{10} + 117n_8 - 324n_6 + 351n_4 - 162n_2 + 27],
\end{aligned} \tag{60}$$

where $\xi_{ijk} \equiv 1/\sqrt{i!j!k!}$ and $n_{2k} \equiv E_\phi[x^{2k}] = \frac{(2k)!}{2^k k!}$. 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 (8).

Proof of Proposition 3 (ES). Consider $\xi_{ijk} \equiv 1/\sqrt{i!j!k!}$, the equation in (55) and $B_k(x)$ in (57), then the expressions of $\Gamma_{ijk}(\cdot)$ in (12) are easily computed as

$$\begin{aligned}
\Gamma_{100}(x) &= B_1(x), \\
\Gamma_{130}(x) &= \xi_{130} [B_4(x) - 3B_2(x)], \\
\Gamma_{140}(x) &= \xi_{140} [B_5(x) - 6B_3(x) + 3B_1(x)], \\
\Gamma_{133}(x) &= \xi_{133} [B_7(x) - 6B_5(x) + 9B_3(x)], \\
\Gamma_{134}(x) &= \xi_{134} [B_8(x) - 9B_6(x) + 21B_4(x) - 9B_2(x)], \\
\Gamma_{144}(x) &= \xi_{144} [B_9(x) - 12B_7(x) + 42B_5(x) - 36B_3(x) + 9B_1(x)].
\end{aligned} \tag{61}$$

Proof of Corollary 2 (LPM). Consider $\xi_{ijk} \equiv 1/\sqrt{i!j!k!}$, the expression in (55) and $B_k(x)$ in (57), then $\Gamma_{ijk}(\cdot)$ in (17) are obtained as

$$\begin{aligned}
\Gamma_{200}(x) &= \xi_{200} [B_2(x) - B_0(x)], \\
\Gamma_{230}(x) &= \xi_{230} [B_5(x) - 4B_3(x) + 3B_1(x)], \\
\Gamma_{240}(x) &= \xi_{240} [B_6(x) - 7B_4(x) + 9B_2(x) - 3B_0(x)], \\
\Gamma_{233}(x) &= \xi_{233} [B_8(x) - 7B_6(x) + 15B_4(x) - 9B_2(x)], \\
\Gamma_{234}(x) &= \xi_{234} [B_9(x) - 10B_7(x) + 30B_5(x) - 30B_3(x) + 9B_1(x)], \\
\Gamma_{244}(x) &= \xi_{244} [B_{10}(x) - 13B_8(x) + 54B_6(x) - 78B_4(x) + 45B_2(x) - 9B_0(x)].
\end{aligned} \tag{62}$$

Proof of Proposition 4 (general expression of σ_ε^2). Following He and Teräsvirta (1999), we can rewrite (24) as $\sigma_t = \alpha_0 + \sigma_{t-1}c_{t-1}$ where $c_t = \beta + \alpha_1^+ z_t^+ - \alpha_1^- z_t^-$ and so, $\sigma_t^2 = \alpha_0^2 + c_{t-1}^2 \sigma_{t-1}^2 + 2\alpha_0 c_{t-1} \sigma_{t-1}$. By taking expectations in both sides of this equation, it is verified that $E(c_t^2) = E(c_{t-1}^2)$, $E(c_{t-1} \sigma_{t-1}) = E(c_t) E(\sigma_t)$. If we assume the stationarity condition that $E(\sigma_t^k) = E(\sigma_{t-1}^k)$ with $k = 1, 2$, then we obtain $E(\sigma_t) = \alpha_0 [1 - E(c_t)]^{-1}$ and $E(\sigma_t^2) = \alpha_0^2 [1 - E(c_t)]^{-1} [1 - E(c_t^2)]^{-1}$. Finally, to achieve the closed-form formula of $E(\sigma_t^2)$ in (25), we need to obtain $E(c_t)$ and $E(c_t^2)$. We take the following steps. First, it is verified that $E(z_t^k) = E[(z_t^+)^k] + E[(z_t^-)^k]$ with $k \in \mathbb{N}$, then $E(z_t^+) = -E(z_t^-)$ since $E(z_t) = 0$ and $E[(z_t^+)^2] = 1 - E[(z_t^-)^2]$ since $E(z_t^2) = 1$. Second, it can be easily obtain $E(c_t) = \beta - (\alpha_1^- + \alpha_1^+) E(z_t^-)$ and denoted as ϖ_1 in (26). Third, given some previous results, we can express $E(c_t^2)$ in terms of $E[(z_t^-)^k]$ for $k = 1, 2$ and denoted as ϖ_2 in (27).

Proof of Corollary 5 (expression of σ_ε^2 under $N(0, 1)$). Given the previous results in Proposition 4, we can easily obtain σ_ε^2 in (25) when $z_t \sim iid N(0, 1)$. We only have to compute $E[(z_t^-)^k]$ for $k = 1, 2$ and plug into the equations (26) and (27). First, it is verified that $E(z_t^-) = B_1(0) = -\phi(0)$ from Proposition 3, then $\varpi_1 = \beta + (\alpha_1^- + \alpha_1^+) \phi(0)$ and we obtain (28). Second, consider both $x^2 = \sqrt{2}H_2(x) + 1$ and Corollary 2, then $E[(z_t^-)^2] = \sqrt{2}\Gamma_{200}(0) + B_0(0) = B_2(0)$ where $B_2(0) = 1/2$ by using (57) and we finally obtain (29).

Proof of Corollary 6 (expression of σ_ε^2 under standardized TGC). Similar to Corollary 5, we proceed to compute $E[(z_t^-)^k]$ for $k = 1, 2$. Consider the general expression given by

$$E[(z_t^-)^k] = \int_{-\infty}^{-a/b} (a + bx_t)^k q(x_t, \boldsymbol{\theta}) dx_t = \sum_{j=0}^k \binom{k}{j} a^{k-j} b^j \xi_j^*(-a/b, \boldsymbol{\theta}), \quad (63)$$

where $k \in \mathbb{N}$, $\xi_j^*(u, \boldsymbol{\theta}) = \int_{-\infty}^u x^j q(x, \boldsymbol{\theta}) dx$ and $q(\cdot)$ is the pdf of TGC in (5). First, we set $k = 1$ in (63), then

$$E(z_t^-) = \int_{-\infty}^{-a/b} (a + bx_t) q(x_t, \boldsymbol{\theta}) dx_t = a\xi_0^*(-a/b, \boldsymbol{\theta}) + b\xi_1^*(-a/b, \boldsymbol{\theta}),$$

such that $\xi_0^*(-a/b, \boldsymbol{\theta}) = Q(-a/b, \boldsymbol{\theta})$ with $Q(\cdot)$ is the cdf of TGC in (7), $\xi_1^*(\cdot)$ can be rewritten as $\xi_1^*(u, \boldsymbol{\theta}) = \int_{-\infty}^u H_1(x) q(x, \boldsymbol{\theta}) dx = \lambda\Psi_1(u, \boldsymbol{\theta})$ with $\Psi_1(\cdot)$ in (32) where $\Gamma_{1ij}(\cdot)$ is obtained in Proposition 3. Second, for $k = 2$ in (63), then

$$E[(z_t^-)^2] = \int_{-\infty}^{-a/b} (a + bx_t)^2 q(x_t, \boldsymbol{\theta}) dx_t = a^2\xi_0^*(-a/b, \boldsymbol{\theta}) + 2ab\xi_1^*(-a/b, \boldsymbol{\theta}) + b^2\xi_2^*(-a/b, \boldsymbol{\theta}),$$

such that $\xi_2^*(\cdot)$ can be rewritten as $\xi_2^*(u, \boldsymbol{\theta}) = \int_{-\infty}^u (\sqrt{2}H_2(x) + 1) q(x, \boldsymbol{\theta}) dx = \sqrt{2} \int_{-\infty}^u H_2(x) q(x, \boldsymbol{\theta}) dx + Q(-a/b, \boldsymbol{\theta})$. Finally, $\int_{-\infty}^u H_2(x) q(x, \boldsymbol{\theta}) dx = \lambda\Psi_2(u, \boldsymbol{\theta})$ with $\Psi_2(\cdot)$ in (32) where $\Gamma_{2ij}(\cdot)$ is obtained in Corollary 2.

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] Alizadeh, A.H., Gabrielsen, A., 2013. Dynamics of credit spread moments of European corporate bond indexes. *Journal of Banking & Finance* 37(8), 3125-3144.
- [2] Angelidis, T., Degiannakis, S., 2007. Backtesting VaR models: a two-stage procedure. *Journal of Risk Model Validation* 1(2), 27-48.
- [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 hyperpolynomial 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] 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.
- [11] Christoffersen, P.F., Diebold, F.X., 2006. Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science* 52(8), 1273-1287.
- [12] Corrado, C., 2007. The hidden martingale restriction in Gram-Charlier option prices. *Journal of Futures Markets* 27(6), 517-534.
- [13] 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.
- [14] Davidson, R., MacKinnon, J.G., 1998. Graphical Methods for Investigating the Size and Power of Hypothesis Tests. *The Manchester School* 66, 1-26.
- [15] Del Brio, E.B., Mora-Valencia, A., Perote, J., 2017. Risk quantification for commodity ETFs: Backtesting value-at-risk and expected shortfall. *International Review of Financial Analysis*, forthcoming.
- [16] 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.
- [17] Draper, N., Tierney, D., 1972. Regions of positive and unimodal series expansion of the Edgeworth and Gram-Charlier approximation. *Biometrika* 59, 463-465.
- [18] Du, Z., and Escanciano, J. C., 2017. Backtesting expected shortfall: accounting for tail risk. *Management Science* 63 (4), 940-958.
- [19] Escanciano, J. C., Olmo, J., 2010. Backtesting parametric value-at-risk with estimation risk. *Journal of Business and Economic Statistics* 28 (1), 36-51.
- [20] Feunou, B., Jahan-Parvar, M.R., Tédongap, R., 2016. Which parametric model for conditional skewness?. *European Journal of Finance* 22(13), 1237-1271.

- [21] Fishburn, P.C., 1977. Mean-risk analysis with risk associated with below-target returns. *American Economic Review* 67(2), 116-126.
- [22] Francq, C. and Zakoïan, J.M. 2010. *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.
- [23] 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.
- [24] Gallant, A.R., Nychka, D.W., 1987. Semiparametric maximum likelihood estimation. *Econometrica* 55, 363-390.
- [25] Gallant, A.R., Tauchen, G., 1989. Semiparametric estimation of conditionally constrained heterogeneous processes: asset pricing implications. *Econometrica* 57, 1091-1120.
- [26] Hansen, B.E., 1994. Autoregressive conditional density estimation. *International Economic Review* 35(3), 705-730.
- [27] He, C., Teräsvirta, T., 1999. Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173-192.
- [28] Jondeau, E., Rockinger, M., 2001. Gram-Charlier densities. *Journal of Economic Dynamics and Control* 25(10), 1457-1483.
- [29] Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness and kurtosis: Existence, persistence, and comovements. *Journal of Economic Dynamics and Control* 27, 1699-737.
- [30] Kerkhof, J., Melenberg, B., 2004. Backtesting for risk-based regulatory capital. *Journal of Banking & Finance* 28, 1845-1865.
- [31] Kräussl, R., Lehnert, T., Senulyté, S., 2016. Euro crash risk. *Journal of Empirical Finance* 38, 417-428.
- [32] Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 2, 174-184.
- [33] 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.
- [34] Lopez, J.A., 1999. Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review* 2, 3-17.
- [35] León, A., Rubio, G., Serna, G., 2005. Autoregressive conditional volatility, skewness and kurtosis. *Quarterly Review of Economics and Finance* 45, 599-618.
- [36] 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.
- [37] León, A., Moreno, M., 2017. One-sided performance measures under Gram-Charlier distributions. *Journal of Banking and Finance* 74, 38-50.
- [38] León, A., Níguez, T.M., 2018. Time-varying semi-nonparametric distribution and portfolio performance, available at SSRN: <https://ssrn.com/abstract=3217148>.
- [39] Liqueur, B., Nazarathy, Y., 2015. A dynamic view to moment matching of truncated distributions. *Statistics and Probability Letters* 104, 87-93.
- [40] Liu, X., Luger, R., 2015. Unfolded GARCH models. *Journal of Economic Dynamics and Control* 58, 186-217.
- [41] Lönnbark, C., 2016. Approximation methods for multiple period Value at Risk and Expected Shortfall prediction. *Quantitative Finance* 16(6), 947-968.

- [42] 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.
- [43] 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.
- [44] Polanski, A., Stoja, E., 2010. Incorporating higher moments into value-at-risk forecasting. *Journal of Forecasting*, 29(6), 523-535.
- [45] 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.
- [46] Skoulakis, G. 2019. Simulating from polynomial-normal distributions. *Communications in Statistics-Simulation and Computation* 48(2), 472-477.
- [47] 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.
- [48] Wu, X., Xia, M., Zhang, H., 2019. Forecasting VaR using realized EGARCH model with skewness and kurtosis. *Finance Research Letters*, forthcoming.
- [49] Zakoïan, J.M., 1994. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* 18 (5), 931-955.
- [50] Zhu, D. and Galbraith, J.W., 2011. Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions. *Journal of Empirical Finance* 18, 765-778.
- [51] 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.
- [52] 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.