Moments expansion densities for quantifying financial risk

Ñíguez, T.M. and Perote, J.

Moments expansion densities for quantifying financial risk

Trino-Manuel Ñíguez

Department of Economics and Quantitative Methods, Westminster Business School, University of Westminster, London, UK

Javier Perote*

Department of Economics and IME, University of Salamanca, Salamanca, Spain

Abstract

We propose a novel semi-nonparametric distribution that is feasibly parameterized to represent the non-Gaussianities of the asset return distributions. Our Moments Expansion (ME) density presents gains in simplicity attributable to its innovative polynomials, which are defined by the difference between the n*th* power of the random variable and the n*th* moment of the density used as the basis. We show that the Gram-Charlier distribution is a particular case of the ME-type of densities. The latter being more tractable and easier to implement when quadratic transformations are used to ensure positiveness. In an empirical application to asset returns, the ME model outperforms both standard and non-Gaussian GARCH models along several risk forecasting dimensions.

Keywords: GARCH; Gram-Charlier series; High-order moments; non-Gaussian distributions; Semi-nonparametric methods; Value-at-Risk.

JEL classification codes: C16, C53, G12.

*Corresponding author: Facultad de Economía y Empresa (Edif. FES), Campus de Miguel de Unamuno, Universidad de Salamanca, 37007 Salamanca, Spain; Tel: +34 923 294400 (ext. 3512); E-mail: perote@usal.es

1 Introduction

The low performance of many traditional methods for financial risk modeling and forecasting during the recent credit crunch has highlighted the deficiencies of standard models for capturing the high-order moments and salient stylized regularities of the asset return distributions (Cont, 2001). Generalized autoregressive conditional heteroscedasticity (GARCH) models (Engle (1982) and Bollerslev (1986)), comprehensively reviewed in Terasvirta (2009), have been extended to account for non-Gaussianities. The alternative densities that have been proposed in previous studies for that purpose include: (i) parametric probability density functions (pdfs henceforth), for instance, the standardized Student's t (Bollerslev, 1987), the GED (Nelson, 1991), the skewed t (Hansen, 1994), the normal inverse Gaussian (NIG) (Jensen and Lunde, 2001), mixtures of normals (Alexander and Lazar, 2006), BEGE combination of gammas distribution (Bekaert et al., 2014), Variance-Gamma (Göncü and Yang, 2016); (ii) non-parametric pdfs (Engle and Gonzalez-Rivera, 1991); and (iii) semi-nonparametric (SNP) densities based on Gram-Charlier (GC) series expansions (Charlier, 1905), introduced in econometrics by Sargan (1976) and developed by authors such as, Gallant and Nychka (1987), Mauleón and Perote (2000), and León et al. (2009), among others.

The current global financial market scenario requires that any assumed density for the shocks in GARCH-type models is flexible enough to account for the leptokurtosis and multimodality of the empirical asset return distributions. In this respect, it is well known that heavy-tailed parametric pdfs overestimate the frequencies for mid-to-lower quantiles, as they try to capture the accumulation of observations in the distribution tail through a monotonic decay. On the other hand, non-parametric densities are more flexible to fit jumps in the distribution tail but they require very large data sets in order to achieve a reasonable degree of precision. Alternatively, SNP pdfs, characterized by its flexibility to represent any frequency function at any degree of accuracy (Cramér 1925), are capable of fitting the wavy shape of the return distributions tail, exacerbated by periods of high financial instability. This fact has awakened a renewed interest for SNP methods and their applications for measuring financial risk– see, e.g., Huang et al. (2014); Lin et al. (2015); Ñíguez and Perote (2016) or León and Moreno (2017). SNP methods, however, also have well-known drawbacks:

(i) Truncated SNP functions (i.e. finite series expansions) are not really pdfs since they may yield negative values (Barton and Dennis, 1952). This issue has been addressed through either parametric constraints à la Jondeau and Rockinger (2001), or density reformulations based on the methodology of Gallant and Nychka (1987) and Gallant and Tauchen (1988) (GNT hereafter).

(ii) Complexity: (a) The direct interpretation of moments in terms of the density parameters is lost when GNT transformations are applied; (b) the characterizations of the density in terms of either the cumulative distribution function (cdf) or the moments generating function (mgf) for GNT-GC pdfs are difficult to obtain; and (c) maximum likelihood (ML) suboptimization is likely to occur.

(iii) SNP pdfs are sensitive to choices in the number of expansion terms.

In this study we present a SNP pdf that, preserving the flexibility typical of GC pdfs, allows to addressing the aforementioned complexities. To do so, we introduce an original series expansion, whose terms are defined as the difference between the n*th* power of the variable and the n*th* moment of the parametric density used as the basis of the expansion. Our Moments Expansion density (ME henceforth) presents gains in simplicity that ease both its theoretically analysis, and practical implementation to model high-order moments and risk measures. We show that the ME is a general family of distributions that nests the GC when the Gaussian density is taken as the basis and GNT transformations are not implemented.

In an empirical application to asset returns, we test the applicability of our model in terms of its relative performance for multiperiod density forecasting as well as for predicting overall measures of market risk, such as, volatility and Value-at-Risk (VaR). The alternative distributions we consider are: Gaussian (used as framework); standardized Student's t; symmetric and skewed GC; and NIG.[1] The models ability for forecasting the conditional variance is measured through ranking-robust loss functions for imperfect volatility proxies (Patton, 2010). VaR forecasting accuracy is assessed through a battery of tests, namely, the magnitude of exceptions statistic (López, 1999), likelihood ratio (LR) tests (Christoffersen, 1998), and the HIT test of Engle and Manganelli (2004). For multiperiod density forecasting we follow the methodology proposed by Maheu and McCurdy (2011). Our results show that Gaussian-ME models are capable of outperforming both standard and non-Gaussian GARCH models for conditional density forecasting, as well as, along several dimensions of market risk forecasting.

The remainder of the paper is organized as follows. In Section 2, we define the ME pdf and analyze its statistical properties. Section 3 provides a comparative analysis of the ME model for forecasting density and financial risk. In Section 4, we summarize our conclusions. All of the proofs are provided in the Appendix.

[1] Other interesting parametric densities to consider include the skewed Student's t; see Ergen (2014) for a recent study of this density for VaR forecasting.

2 The ME density

In this section, we define the ME pdf and study its characterizations and statistical properties in relation to the GC, which we use as framework.

Let a random variable x be GC (Type A) distributed with pdf given by,

$$\pi(x, \mathbf{d}_n) = \left(1 + \sum_{s=1}^{n} d_s H_s(x)\right) \phi(x), \tag{1}$$

where $\phi(\cdot)$ stands for the standard Normal pdf, $H_s(\cdot)$ denotes the HP of order s, and $d_n = (d_1, d_2, \ldots, d_n)' \in \mathbb{R}^n$ with n being the truncation order of the expansion. The HPs, which can be defined as in equation (2), form an orthogonal basis with respect to $\phi(x)$, which is the grounds for $\pi(x, d_n)$ to integrate up to one.

$$H_s(x) = \frac{(-1)^s}{\phi(x)} \frac{d^s \phi(x)}{dx^s}. \tag{2}$$

In general, SNP distributions can be defined for any continuous and differentiable parametric pdf. However, GC series expansions of non-Normal pdfs have rarely been considered as they deliver rather complex specifications, in particular for GNT-type of reformulations.[2] We show that the ME, defined below, is an alternative to the GC expansion, the former exhibiting advantages in simplicity without compromising flexibility and accuracy.

Definition. A ME of a parametric distribution $g(\cdot)$ has a pdf given by,

$$f(x, \boldsymbol{\gamma}_n) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s(x)\right) g(x) = \Lambda_s(x) g(x), \tag{3}$$

where γ_s denotes the sth parameter and $\{\Psi_s(x)\}_{s=1}^{n}$ is a polynomial sequence of the form

$$\Psi_s(x) = x^s - \mu_s, \tag{4}$$

where μ_s denotes the sth raw moment of $g(x)$.

By definition, the ME pdf requires that the distribution used as the basis of the expansion has finite moments up to the truncation order, but no other condition is necessary. By construction, parameter γ_s captures the weight assigned to the deviation of the sth order moment of $f(\cdot)$ with respect to its counterpart in $g(x)$. Two main differences between GC and ME expansions are worth mentioning:

[2] The expansions of the Poisson, Gamma and Beta distributions are known as Gram-Charlier Type B, Laguerre and Jacobi series, respectively.

(i) The ME polynomials, $\Psi_s(x)$, unlike the HPs, are not orthogonal with respect to the distribution $g(x)$, i.e. $\int \Psi_s(x)\Psi_j(x)g(x)dx \neq 0$, $s \neq j$. However, this is not an obstacle to guarantee up to one integration, we shall discuss this property in the next section. On the contrary, $\Psi_s(x)$ has an important advantage in terms of simplicity compared to $H_s(x)$; a feature that is of particular interest when GNT transformations are used to ensure positiveness.

(ii) The ME admits any distribution $g(x)$ as the basis of the expansion, whilst the GC has only been defined for the Gaussian distribution, $\phi(x)$. Note that although the GC approach could be applied to other continuous and differentiable pdfs these extensions would be at the cost of a significant increase in complexity and intractability.[3]

The ME pdf is symmetric when the $g(x)$ is symmetric and $\gamma_s = 0$ $\forall s \geq 3$, s odd. The shape of the ME family of densities with elliptical basis distributions resembles the well-known "Gaussian bell" for a wide range of the parametric space, but its tails may be thicker and wavy. When the parameters γ_s are all 0, the ME nests the density used as the basis. Figure 1 illustrates the density of a Gaussian-ME (GME hereafter) distribution. The plot shows how the GME shape responds to changes in the values of the parameters allowing for heavy tails and multimodality. It also illustrates the convergence of the GME to the Gaussian distribution as γ_s tend to 0.

[Insert Figure 1]

Figure 2 provides the kurtosis range of GNT GME densities truncated at $n = 8$. We represent two- and one-parameter symmetric densities reformulated for positiveness through the GNT transformation presented in Proposition 5 below.[4] The graphs illustrate how the range of kurtosis expands as the truncation order increases. Note that the maximum Range=[1.12,5.892] is obtained by the one-parameter GNT GME with $\gamma_8 \neq 0$, and that adding lower order parameters do not seem to broaden the kurtosis range.

[Insert Figure 2]

[3] The performance of the ME when non-Gaussian densities are used as the basis (in particular the expansion of a Student's t distribution) will be investigated in further research. Whilst in the present paper, we mainly focus on the analysis of the advantages of the ME for defining positive densities when the Gaussian distribution is used as the basis.

[4] The ranges of skweness and kurtosis of skewed GME coincide with those of the GC densities in Leon *et al.* (2005, 2009).

2.1 ME statistical properties

In this section we analyze the ME properties (Propositions 1-9). All of the proofs (except that of Proposition 4 as it provides a straightforward result) are presented in the Appendix. Without loss of generality in the analysis that follows, we consider a standardized $g(x)$, i.e. with mean 0 and variance 1.

Proposition 1. The ME distribution defined in equation (3) integrates up to one: $\int f(x,\gamma_n)dx = 1$.

Proposition 2. The ME is positive in a wide range of its parameter space, particularly $f(x,\gamma_n) \geq 0 \; \forall \gamma_n \in \mathbb{R}^n$ such that

$$\begin{aligned} 0 \leq \gamma_s \leq \frac{1}{n\mu_s} \qquad & \forall s \text{ even} \\ \gamma_s = 0 \qquad & \forall s \text{ odd} \end{aligned} \tag{5}$$

These constraints are obtained straightforwardly, so the ME positivity issue could be easily addressed this way, at least for symmetric expansions (i.e. $\gamma_s = 0 \; \forall s$ odd).[5] Unlike the non-negativity (NN hereafter) constraints of Jondeau and Rockinger (2001), our NN conditions are valid for any truncation order, and imply squared, instead of elliptical, positivity sets. NN constraints and GNT transformations are alternative solutions to ensure the positivity of SNP pdfs, and both present shortcomings. Namely, NN conditions may reduce the goodness-of-fit of the density if the feasible (for positivity) parametric space shrinks too much. Besides, GNT transformations may lead to non-linearities between density moments and parameters, which hinders the interpretation of the parameter estimates in terms of moments. This last feature will be discussed later on.

Proposition 3. The raw moments of the ME distribution, $f(x,\gamma_n)$, are given by:

$$m_i = E[x^i] = \mu_i + \sum_{s=1}^{n} \gamma_s(\mu_{s+i} - \mu_s\mu_i), \;\; \forall i \in \mathbb{N}. \tag{6}$$

Proposition 3 shows that the non-central moments of the ME pdf are obtained from the moments of the density used as the basis, and are linearly related to the ME parameters. Furthermore, the relation in equation (6) reveals that the odd/even moments depend only on the odd/even parameters. Therefore if $g(x)$ is symmetric and all odd parameters are zero

[5] Note that, as skewness can be achieved by using a skewed $g(x)$, the use of symmetric expansions does not represent a serious shortcoming.

then the ME has zero mean and is symmetric. This is an interesting property that simplifies parameter estimation and interpretation. In particular, the even parameters capture the wavy and thick-tailed behavior of the density, whilst the odd parameters incorporate different type of asymmetries.

The ME distribution can alternatively be expressed in terms of its own non-central moments as

$$f(x, \mathbf{M}_n) = \left(1 + \sum_{s=1}^{n} (m_s - \mu_s)\theta_s(x)\right) g(x), \tag{7}$$

where $M_n = (m_1, m_2, \ldots, m_n)' \in \mathbb{R}^n$ with m_i satisfying the relation in equation (6) and $\theta_s(x)$ are the polynomials such that $f(x, M_n) = f(x, \gamma_n)$ – see the example in equation (18). This type of re-parameterization is useful for easing (i) the interpretation of the estimates in terms of the distribution's skewness and kurtosis; and (ii) the comparison with the GC distribution.

Proposition 4. Let $m_i^* = E[(x - m_1)^i]$ $\forall i \in \mathbb{N}$, denote the central moments of $f(x, \gamma_n)$, then a standardized ME distribution, i.e. with zero mean and variance one, denoted as $f^*(z, \gamma_n)$, can be expressed in terms the parameters γ_s as:

$$f^*(z, \boldsymbol{\gamma}_n) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s\left(m_2^{*1/2} z + m_1\right)\right) g\left(m_2^{*1/2} z + m_1\right) m_2^{*1/2}, \tag{8}$$

Note that Proposition 4 is just the result of the linear transformation, $z = (x - m_1)/m_2^{*1/2}$. This expression is useful for introducing GARCH dynamics through $u_t = h_t^{\frac{1}{2}} x_t$, where h_t is the variance of the conditional distribution of u_t.

Proposition 5. A positive ME distribution, $F(x, \boldsymbol{\gamma}_n)$, can be obtained through a GNT transformation, for instance, by squaring each element of the weighting sum as:

$$F(x, \boldsymbol{\gamma}_n) = \frac{1}{W_n}\left(1 + \sum_{s=1}^{n} \gamma_s^2 \Psi_s(x)^2\right) g(x), \tag{9}$$

where W_n is the constant that guarantees that $F(x, \gamma_n)$ integrates up to one,

$$W_n = \int \left(1 + \sum_{s=1}^{n} \gamma_s^2 \Psi_s(x)^2\right) g(x) dx = 1 + \sum_{s=1}^{n} \gamma_s^2 (\mu_{2s} - \mu_s^2). \tag{10}$$

Proposition 5 shows that the ME pdf admits GNT transformations to ensure positivity. It is worth noting that the transformation we have presented in this proposition is just one possibility, among others of GNT-type, that ME densities admit to ensure positivity.

For example, other GNT reformulations such as the one below, i.e., squaring the whole weighting polynomial sum, could also be implemented but would produce more complex density specifications,

$$\mathcal{F}(x,\boldsymbol{\gamma}_n)=\frac{1}{\omega_n}\Lambda_s(x)^2 g(x), \tag{11}$$

where $\omega_n \equiv \int \Lambda_s(x)^2 g(x)dx$.

$F(x,\boldsymbol{\gamma}_n)$ yields a symmetric pdf provided that $g(x)$ is symmetric, while the positivity transformation in $\mathcal{F}(x,\boldsymbol{\gamma}_n)$ allows for skewness even when $g(x)$ is symmetric, but results in a more complex specification as the expansion of $\Lambda_s(x)$ includes squared polynomials, $\Psi_s(x)^2$, as well as cross products, $\Psi_i(x)\Psi_j(x)$ $i \neq j$. This feature makes easier to relate the density's moments to its parameters, which simplifies its tractability and applicability. For instance, the variance and kurtosis of a symmetric four-order ME, $F(x,\boldsymbol{\gamma}_4)$, depend only on the squared parameters, γ_2^2 and γ_4^2, which saves cross-products, $\gamma_2 \cdot \gamma_4$, non-linearities in the moments expressions. In what follows, we restrict our analysis to the GNT transformation in $F(x,\boldsymbol{\gamma}_n)$, denoted as GNT-ME.[6]

Proposition 6. The raw moments of the positive ME distribution in Proposition 5, $F(x,\boldsymbol{\gamma}_n)$, can be expressed in terms of the moments of the expanded density, $g(x)$, and the squared density parameters as:

$$\widetilde{m}_i = E[x^i] = \frac{1}{W_n}\left[\mu_i + \sum_{s=1}^{n}\gamma_s^2\left[\mu_{2s+i}+\mu_s(\mu_s\mu_i - 2\mu_{s+i})\right]\right],\ \forall i=1,2,..., \tag{12}$$

The relation between the GNT-ME moments and parameters provided by Proposition 6, is slightly more complex than that of equation (6), but is still much simpler than the corresponding one to the GNT-GC case. The property in this proposition is interesting because several reasons: (i) it eases the interpretation of the GNT-ME's parameters in relation to those of the GNT-GC; (ii) it simplifies estimation; and (iii) it allows for a direct implementation of dynamics of GARCH-type for high-order moments.

2.2 The Gaussian ME density

In this section we analyze the properties of the GME, i.e. the ME with the Gaussian pdf, $\phi(x)$, used as the basis of the expansion. We will show that in the absence of GNT transformations, the GME inherits the good asymptotic properties of the GC approximation, thus being a good alternative for modeling and forecasting high-order moments.

[6] See Ñíguez and Perote (2012) for a comprehensive analysis of these two types of GNT transformations applied to GC pdfs. Comparisons of multivariate related extensions can be found in Del Brio et al. (2009; 2011).

Let us denote the GME pdf as

$$f_N(x, \boldsymbol{\gamma}_n) = \left(1 + \sum_{s=1}^{n} \gamma_s (x^s - \mu_s^+)\right) \phi(x), \tag{13}$$

where μ_s^+ denotes the s-th order raw moment of $\phi(x)$.

2.2.1 Relation between GME and GC

Proposition 7. The GC can be expressed as the GME pdf for certain restrictions on their parameters $\mathbf{d}_n$ and $\boldsymbol{\gamma}_n$ and provided that they both are not GNT transformed.

$$\pi(x, \mathbf{d}_n) = \left(1 + \sum_{s=1}^{n} d_s H_s(x)\right) \phi(x) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s(x)\right) \phi(x) = f_N(x, \boldsymbol{\gamma}_n). \tag{14}$$

Proposition 7 shows that the GC is a reparameterization of the GME and a particular case of the ME for $g(x) = \phi(x)$. As a result, the GME is an asymptotically valid SNP method to approximate any pdf. An appealing feature of the GME, with respect to the GC, is its practicality to address the positivity problem, as GNT transformations of the GME are much simpler than those of the GC. This simplicity arises from the fact that GNT-GME only involves squares of binomials instead of polynomials whose orders increase with the expansion terms. This feature is illustrated in the example below, where we review the GME properties in relation to the GC expansion.

2.2.2 Analytical example

We discuss the properties of the GME presented so far using an example of a GME pdf truncated at order four ($n = 4$), which is the most widely used case in empirical applications of GC densities. Let us consider a GC truncated at $n = 4$,

$$\pi(x, \mathbf{d}_4) = \left[1 + d_1 x + d_2(x^2 - 1) + d_3(x^3 - 3x) + d_4(x^4 - 6x + 3)\right] \phi(x). \tag{15}$$

By rearranging terms we obtain a four-order GME pdf:

$$\begin{aligned} \pi(x, \mathbf{d}_4) &= \left[1 + (d_1 - 3d_3)x + (d_2 - 6d_4)(x^2 - 1) + d_3 x^3 + d_4(x^4 - 3)\right] \phi(x) \\ &= \left[1 + \gamma_1 x + \gamma_2(x^2 - 1) + \gamma_3 x^3 + \gamma_4(x^4 - 3)\right] \phi(x) = f_N(x, \boldsymbol{\gamma}_4). \end{aligned} \tag{16}$$

The first four raw moments of the pdf in (16) are: $m_1 = \gamma_1 + 3\gamma_3$, $m_2 = 2\gamma_2 + 12\gamma_4$, $m_3 = 3\gamma_1 + 15\gamma_3$, and $m_4 = 12\gamma_2 + 94\gamma_1$.

$\pi(x,\mathbf{d}_4)$ can alternatively be re-written in terms of its first four moments as,

$$f_N(x,\mathbf{M}_4)=\left[1+m_1\theta_1(x)+(m_2-1)\theta_2(x)+m_3\theta_3(x)+(m_4-3)\theta_4(x)\right]\phi(x), \tag{17}$$

where

$$\theta_i(x)=\begin{cases} x(5-x^2)/2, & if\ i=1,\\ 2(x^2-1)-(x^4-3)/4, & if\ i=2,\\ x(x^2+3)/6, & if\ i=3,\\ (x^4-3)/24-(x^2-1)/4, & if\ i=4.\end{cases} \tag{18}$$

A standardized $f_N(x,\boldsymbol{\gamma}_4)$ (with $m_1=0$ and $m_2=1$) can be obtained by imposing the restrictions: $\gamma_1=-3\gamma_3$, $\gamma_2=-6\gamma_4$ (or in terms of $\mathbf{d}_4$: $d_1=d_2=0$). The resulting pdf can conveniently be expressed in terms of the skewness ($sk=m_3^*/m_2^{*3/2}$) and kurtosis ($ku=m_4^*/m_2^{*2}$) as:[7]

$$\widetilde{f}_N(x,\mathbf{M}_4)=\left[1+\frac{sk}{3!}(x^3-3x)+\frac{ku-3}{4!}(x^4-6x^2+3)\right]\phi(x). \tag{19}$$

A symmetric case of $f_N(x,\boldsymbol{\gamma}_4)$ may be obtained by considering $\gamma_1=\gamma_3=0$:

$$\widehat{f}_N(x,\boldsymbol{\gamma}_4)=\left[1+\gamma_2(x^2-1)+\gamma_4(x^4-3)\right]\phi(x). \tag{20}$$

For this distribution $0\leq\gamma_2\leq 0.25$ and $0\leq\gamma_4\leq 0.083$ are sufficient NN conditions for the density in equation (20).

Alternatively, a well-defined GME can be obtained by a GNT transformation as:

$$\widehat{F}_N(x,\boldsymbol{\gamma}_4)=\frac{1}{W^*}\left[1+\gamma_2^2(x^2-1)^2+\gamma_4^2(x^4-3)^2\right]\phi(x), \tag{21}$$

where $W^*=1+2\gamma_2^2+96\gamma_4^2$. The variance of $\widehat{F}_N(x,\boldsymbol{\gamma}_4)$ is $m_2=\left(1+10\gamma_2^2+864\gamma_4^2\right)/W^*$ and then the pdf of the standardized, symmetric and positive GME truncated at $n=4$ is

$$\widehat{F}_N^*(x,\boldsymbol{\gamma}_4)=\frac{1}{W^*}\left[1+\gamma_2^2(m_2x^2-1)^2+\gamma_4^2(m_2^2x^4-3)^2\right]\phi(m_2^{1/2}x)m_2^{1/2}. \tag{22}$$

[7] This is the traditional expression for the GC density (e.g. León *el al.*, 2005) used for capturing the impact of skewness and kurtosis on returns distribution and its applications to option princing.

2.2.3 GME cumulative distribution functions

In this section we characterize the GME by means of its cdf. We show that the GME's cdf can be easily obtained from the cdf of the normal distribution. Propositions 8 and 9 below present closed forms for the GME's cdf and GNT positive version.

Proposition 8. The cdf of a random GME distributed variable is given by,

$$\begin{aligned}\Pr[x \le a] &= \int_{-\infty}^{a} f_N(x, \boldsymbol{\gamma}_n)dx = \int_{-\infty}^{a} \phi(x)dx \\ &\quad - \sum_{s=1}^{n} \gamma_s \left(a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \xi a^b\right)\phi(a), \quad (23)\end{aligned}$$

where

$$\xi = \begin{cases} (s-1)(s-3)\cdots 2, & \forall s \text{ odd}, \\ (s-1)(s-3)\cdots 3, & \text{otherwise}, \end{cases} \quad (24)$$

and

$$b = \begin{cases} 1, & \forall s \text{ even}, \\ 0, & \text{otherwise}. \end{cases} \quad (25)$$

Proposition 9. The cdf of a variable distributed as a GNT-GME distribution given by $F_N(x, \gamma_n)$ is,

$$\begin{aligned}\Pr[x \le a] = &\int_{-\infty}^{a} F_N(x, \boldsymbol{\gamma}_n)dx = \int_{-\infty}^{a} \phi(x)dx \\ &+ \frac{2}{W_n}\sum_{s=1}^{n} \gamma_s^2 \mu_s^+ \left(a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \zeta a^b\right)\phi(a) \\ &- \frac{1}{W_n}\sum_{s=1}^{n} \gamma_s^2 \left(a^{2s-1} + (2s-1)a^{s-3} + (2s-1)(2s-3)a^{2s-5} + \ldots + \mu_{2s}^+ a^b\right)\phi(a),\end{aligned}$$

(26)

where W_n is the constant in equation (10) for the moments of $\phi(x)$, denoted as μ_s^+, b is the constant defined in equation (25),

$$\zeta = \begin{cases} (s-1)(s-3)\cdots 2, & \forall s \text{ odd}, \\ (s-1)(s-3)\cdots 3a, & \text{otherwise}. \end{cases} \quad (27)$$

2.3 Further extensions

The ME represents a very general class of densities, the GME being the only case that has been analysed so far. The study of other particular cases are out of the scope of the present article but anyhow we mention two natural extensions of the GME that deserve particular analyses.

The first one is the ME of mixtures of Gaussian densities. Particularly, for the purpose of identification we concentrate on the mixture of two normal distributions with zero mean and variance σ_i^2, denoted by $\phi_{\sigma_i}(x)$, $i = 1, 2$, whose pdf is presented in equation (28).

$$f_{MG}(x, \boldsymbol{\gamma}_n) = \left(1 + \sum_{s=1}^{n} \gamma_s (x^s - \alpha_s \mu_s^+)\right) \left[p_1 \phi_{\sigma_1}(x) + p_2 \phi_{\sigma_2}(x)\right] \tag{28}$$

where $\alpha_s = p_1 \sigma_1^s + p_2 \sigma_2^s$, $0 \leq p_i \leq 1$ and $p_1 + p_2 = 1$. It is clear that the GME is a particular case of this density provided that $\sigma_i^2 = 1$ and $p_i = 0.5$, and $\forall i = 1, 2$.

The second is the expansion of the Student's t density with v degrees of freedom, $t_v(x)$, whose pdf is characterized below.

$$f_{ST}(x, \boldsymbol{\gamma}_n) = \left(1 + \sum_{s=1}^{n} \gamma_s (x^s - \varphi_s)\right) t_v(x) \tag{29}$$

where $\varphi_s = v^s \frac{\Gamma\left(\frac{v+1}{2}\right)\Gamma\left(\frac{v-s}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{v}{2}\right)}$ for s even and $\varphi_s = 0$ for s odd, and $\Gamma(\cdot)$ denotes the gamma function. This density also encompasses the GME as v tends to infinity.

The in-depth analysis of the properties of such densities is left for further research.

3 Empirical application

3.1 Data, modeling and in-sample performance

This section provides an analysis of the applicability of the ME through an in- and out-of-sample comparative analysis. We use the data set from a previous study in Ñíguez and Perote (2012). The data are (daily) log returns in percentage, computed from samples of asset prices, $r_t = 100 \log(P_t/P_{t-1})$, of Microsoft sampled over the period 1/16/1989 to 1/15/2009 for a total of $T = 5,219$ observations; the British Pound(£)/US Dollar($) exchange rate (FX £/$) from 1/16/1989 to 1/13/2009, $T = 5,217$; and the S&P 500 index, from 2/15/1988 to 2/16/2004, $T = 4,176$. Two time periods are considered (stable and volatile) to avoid dependence of results on the sample time span. The series show stylized features of financial returns such as: small first-order correlation in the level, volatility clustering, mild skewness and high excess kurtosis.

We consider alternative specifications that have been used in the literature for modeling the density of asset returns. Let the conditional distribution of r_t with respect to the information set available up to time $t-1$ (Ω_{t-1}) be: Gaussian, $u_t|\Omega_{t-1} \sim N(0, h_t)$; standardized Student's t with ν degrees of freedom, $u_t|\Omega_{t-1} \sim t_\nu(0, h_t)$; (symmetric) NIG with steepness parameter $\zeta \geq 0$, $u_t|\Omega_{t-1} \sim NIG(0, h_t)$; GNT GC, $u_t|\Omega_{t-1} \sim GC(0, h_t)$; GNT skewed-GC, $u_t|\Omega_{t-1} \sim sk\text{-}GC(0, h_t)$; GNT (symmetric) GME, $u_t|\Omega_{t-1} \sim GME(0, h_t)$, equation (22).[8],[9] The conditional mean and variance of r_t follow an AR(1) – selected according to the Akaike Information Criterion (AIC) – and AGARCH(1,1) processes (Glosten et al., 1993), equations (30) and (31), respectively,

$$r_t = \delta_0 + \delta_1 r_{t-1} + u_t, \quad u_t = h_t^{\frac{1}{2}} x_t, \tag{30}$$

$$h_t = \varphi_0 + \varphi_1 \left(|u_{t-1}| - \xi u_{t-1}\right)^2 + \varphi_2 h_{t-1}, \tag{31}$$

The estimation procedure is carried out in two steps. We use the first $T - R$ observations as the first in-sample window, and compute $R = 1,000$ out-of-sample forecasts using a rolling window that discards old observations. The models are estimated by quasi-ML (QML), robust covariance estimators are computed using Bollerslev and Wooldridge (BW) (1992) formula. Table 1 presents the estimation results.[10] As a first observation, the optimization of the GME likelihood function was smoothly achieved after starting values were chosen adequately, those values were obtained using the sample moments and equation (12). The estimation of the conditional variance is similar for all models and series; it gathers the typical strong persistence in the second moment of asset returns ($\widehat{\varphi}_1 + \widehat{\varphi}_2$ close to 1) and

[8] We specify GNT SNP models (GC and GME) truncated at $n = 4$. The GNT GME model considers the more simple and symmetric GNT transformation in Ñíguez and Perote (2012), but the GC densities incorporate the GNT reformulation in León et al. (2009). In order to analyze the effect on forecasting performance of capturing the mild skewness of the data empirical distribution, we also consider the GNT sk-GC density in León et al. (2005).

[9] A comparison of current alternative approaches for risk prediction, including historical simulation and variants is provided in Kuester *et al.* (2006), and Bao *et al.* (2006).

[10] The estimation of the conditional mean gathers the small structure in the level of the return series. This estimation is not affected by the density assumption. The parameter estimates $\widehat{\delta}_0$ and $\widehat{\delta}_1$ are not presented in Table 1 for saving space.

shows that the leverage effect is present in the returns volatility (since H_0: $\xi = 0$ is rejected). All parameters ν, ζ, γ_2 and γ_4 are significant confirming that the empirical distribution is leptokurtic. The estimate of γ_3 accounts for the skewness of the distribution and yields a small negative value for all series, that is significant only for the S&P 500 returns. Regarding the goodness-of-fit, all models provide better results than the Gaussian; this is attributable to the fact that the Gaussian distribution is not able to capture the existing leptokurtosis. As a conclusion, our GME model yields an AIC that is lower than the Gaussian and it is in the range of the AICs obtained from the rest of non-Gaussian models.

Figure 3 provides an illustration of the fit and shapes of the GME distribution in relation to the Student's t and Gaussian for the FX £/S return series. All three distributions fit reasonably well the probabilistic mass at the center (variance) of the histogram. However, the plots show that the Gaussian distribution cannot capture the leptokurtosis of the distribution (i.e. underestimates risk), whilst the Student's t overestimate the frequency of the whole tail as it tries to capture the cluster of outliers at the left end of the tail. The GME shows a flexible fit that captures frequencies between the Gaussian and the Student's t, and thus involves more accurate risk measures.

[**Insert Table 1 and Figure 3**]

3.2 Out-of-sample performance

In-sample goodness-of-fit does not provide enough information to conclude on the relative performance of a model, even if different series are considered, as it is based on a single observation. It is well known that a model may provide a better in-sample fit but not perform better for out-sample forecasting. In this section we assess the models' ability to forecast different measures of market risk associated with the second as well as higher-order moments of the return distributions.

The forecast accuracy for the conditional variance, h_t, is measured with respect to the out-of-sample squared residuals, $\{\widehat{u}_t^2\}_{t=T-R+1}^{T}$, by using the mean square forecasting error (MSFE). The $MSFE = \Sigma_{t=T-R+1}^{T}(h_t - \widehat{u}_t^2)^2/R$ is a robust loss function to models ranking when using imperfect volatility proxies; see Patton (2010) and Awartani and Corradi (2005). The significance of the difference between the MSFE from the alternative models is measured by using the Giacomini and White (GW) (2006) test. Table 2 presents the GW test statistic for all pairwise comparisons. The entries are the p-values of the GW test. The results show that SNP models present a higher MSFE with differences tending to be significant only for the GC in relation to the rest of the models. Among the parametric models the ranking is: Student's t, NIG and Gaussian, although differences are very small. It seems that capturing skewness helps the sk-GC in the margin to improve volatility forecasts. As a conclusion, our GME model presents a reasonable performance for forecasting volatility with a MSFE that is not statistically significant with respect to the rest of models used in the analysis.

[**Insert Table 2**]

Next, we test the models performance for forecasting r_t distribution tails. To do so, we compute R 1-step-ahead VaR forecasts for coverage levels $\alpha = \{0.1, 0.05, 0.01\}$, $\widehat{VaR}_{T+1}^{\alpha} = \widehat{r}_{T+1} - \widehat{\kappa}_T \widehat{h}_{T+1}^{1/2}$, where $\widehat{\kappa}_T$ is the α-quantile of the assumed distribution. Following Engle (2001), the VaR forecasts corresponding to the Gaussian model are computed by using the percentile of the empirical distribution of the standardized residuals (hereafter we denote these VaR forecasts as $Normal^E$). The performance of the models is assessed by using the following criteria: unconditional and conditional LR test (Christoffersen, 1998); unconditional coverage, $\widehat{\alpha}$, and magnitude of the exception, M_α, (López 1999); and HIT test (Engle and Manganelli, 2004). Table 3 presents the results of the VaR evaluation criteria for all models and series. A first observation that emerges from this table is that all models yield similar acceptable results for unconditional coverage, α; the Student's t model seeming to systematically underestimate the unconditional probability. The lower values of the magnitude of the exception statistic is performed by the Student's, and the GME model provides a similar performance than the NIG and Normal^E which are overall better than the GC models (except for the FX series). According to the HIT test, all models (except the GC models for S&P 500 and Microsoft) present acceptable VaR forecasts. Respect to LR tests, heavy-tailed models (including the GME) tend to perform better (except the Student's t in this case) the lower the quantile. As a summary, the GME pdf seems to work well for forecasting VaR in relation to the wide range of alternative SNP and parametric pdfs considered in this analysis.

[**Insert Table** 3]

Finally, GME models are evaluated according to their performance for multiperiod density forecasting. To do so, we follow the methodology described in Maheu and McCurdy

(2011), and recently applied in Anatolyev and Petukhov (2016). A model provides better forecasting performance when its average predictive likelihood is higher. The null hypothesis of equal density forecasting performance from two alternative models was tested using the GW test. For our application to FX \$/£ returns series, we computed 1000 out-of-sample forecast for each forecast horizon $k = 1, ..., 60$ days, with random numbers sample size of $10,000$ for the Monte Carlo simulation of returns.[11] We considered two-parameter symmetric GNT GME truncated at $n = 8$ with $\gamma_4 = \gamma_6 = 0$, denoted as GME(2,8); the GME model used for volatility and VaR forecasting; as well as the Student's t and Normal as benchmark models. Figure 4 presents the multiperiod conditional forecast results. Note that all of the average predictive likelihood decrease with k. This is because the accuracy of the forecast conditional on time t decreases with k. The GME(2,8) model performs, significantly better than the Student's t for short horizons, $k = 1, ..., 5$, and worse for larger k, although differences are not significant for the latter. GME(2,8) and Student's t are better than GME(2,4) although differences are not significant for shorter horizons. GME and Student's t models perform better than the Normal and differences become more significant the longer the horizon.

[**Insert Figure** 4]

4 Conclusions

In this study we have proposed a new SNP density that features an original polynomial structure. This density, which we refer to as Moments Expansion (ME), presents the following innovations: (i) it is correctly defined and nests the GC when the Gaussian is used as the basis; (ii) non-negativity parametric constraints are very easy to obtain for symmetric ME densities; (iii) it admits GNT-type transformations to ensure positivity; and (iv) both ME and GNT-ME pdfs can be straightforwardly characterized in terms of their moments and cdfs. Our analysis shows that the ME is simpler and more tractable for theoretical and practical purposes than the GC, offering a feasible way for expanding any parametric pdf – only requiring that the density used as the basis has finite moments up to the truncation

[11] The GME random numbers series where obtained from the inverse of the cdf for $10,000$ random probabilities drawn from U[0,1]. Thus, solving the equation $F(x) = u$, where $F(\cdot)$ is the cdf of the GME (obtained applying equation (A.10)) and u is the generated random number.

order.

The relative performance of a Gaussian-ME has been tested through an empirical application for forecasting the density, conditional variance and VaR of asset returns, considering alternative distributions used in previous studies. Our results show that the GME works well in sample and provides a reasonable performance for forecasting the asset returns density, and second and higher-order moments of return distributions. The ME model is thus a simpler an more accurate alternative to GC densities and other heavy-tailed parametric densities used for risk management.

Finally, future work may include the extension of the ME to the multivariate context as well as its development to expand non-Gaussian pdfs. Of particular interest is the case of the ME for the Student's t (ME-t) distribution because of the following three preliminary reasons: (i) ME-t naturally converges to the GC and the GME as $\nu \to \infty$, which provides a solid framework for the theoretical analysis, (ii) ME-t may enlarge the positivity regions in terms of skewness and kurtosis compared to the GC; and (iii) ME-t entails more flexibility for goodness-of-fit with lower truncation orders thanks to the extra parameter ν. On the other hand, multivariate ME distributions may be useful to deal with the dimensionality problem of large portfolios when implementing conditional co-moments and asymmetric dependence.

Acknowledgements

We thank the editor and two anonymous referees for their valuable comments and suggestions. We also thank seminar participants at the XVII Annual Symposium of the Society of Nonlinear Dynamics and Econometrics, 2nd International Conference on Computational and Financial Econometrics, XXXII Symposium of Economic Analysis, 27th International Symposium on Forecasting, 10th BMRC-DEMS Conference, 15th INFINITI Conference on International Finance, University of Lancaster, Bank of Spain and University of Murcia for comments and suggestions. This work was supported by the Spanish Ministry of Economics and Competitiveness [grants ECO2013-44483-P and ECO2016-75631-P] and Junta de Castilla y León [SA072U16)]. All remaining errors are ours.

Appendix: Proofs of propositions

Proof of Proposition 1. The ME pdf integrates up to one.

$$\int f(x,\boldsymbol{\gamma}_n)dx = \int g\left(x\right)dx + \sum_{s=1}^{n}\gamma_s\int(x^s-\mu_s)g(x)dx = 1+\sum_{s=1}^{n}\gamma_s(\mu_s-\mu_s) = 1.\square \quad \text{(A.1)}$$

Proof of Proposition 2. The conditions in equation (3) are sufficient for the ME density to be positive. We re-write the ME density as,

$$f(x,\boldsymbol{\gamma}_n) = \left(\sum_{s=1}^{n}\gamma_s x^s + k\right)g(x), \quad \text{(A.2)}$$

where $k = 1-\sum_{s=1}^{n}\gamma_s\mu_s$. For this expression it is clear that if $0 \leq \gamma_s \leq \frac{1}{n\mu_s}$ $\forall s$ even, and $\gamma_s = 0$ $\forall s$ odd, then $\sum_{s=1}^{n}\gamma_s x^s + k \geq 0$ and $f(x,\boldsymbol{\gamma}_n) \geq 0.\square$

Proof of Proposition 3. ME moments are linearly related to the density's parameters.

$$\begin{aligned} E[x^i] &= \int x^i f(x,\boldsymbol{\gamma}_n)dx \\ &= \int x^i g\left(x\right)dx + \sum_{s=1}^{n}\gamma_s\int x^i(x^s-\mu_s)g(x)dx \\ &= \mu_i+\sum_{s=1}^{n}\gamma_s(\mu_{s+i}-\mu_s\mu_i).\square \end{aligned} \quad \text{(A.3)}$$

Proof of Proposition 5. The constant W_n that makes the GNT-ME to integrate up to one is

$$\begin{aligned} W_n &= \int\left(1+\sum_{s=1}^{n}\gamma_s^2(x^s-\mu_s)^2\right)g(x)dx \\ &= \int g(x)dx + \sum_{s=1}^{n}\gamma_s^2\left(\int x^{2s}g(x)dx + \mu_s^2\int g(x)dx - 2\mu_s\int x^s g(x)dx\right) \\ &= 1+\sum_{s=1}^{n}\gamma_s^2(\mu_{2s}+\mu_s^2-2\mu_s^2) = 1+\sum_{s=1}^{n}\gamma_s^2(\mu_{2s}-\mu_s^2). \end{aligned} \quad \text{(A.4)}$$

Given W_n it is is clear that $\int F(x,\boldsymbol{\gamma}_n)dx = 1.\square$

Proof of Proposition 6. The GNT-ME moments are functions of its squared parameters.

$$\begin{aligned} E\left[x^i\right] &= \frac{1}{W_n}\int x^i\left(1+\sum_{s=1}^{n}\gamma_s^2(x^s-\mu_s)^2\right)g(x)dx \\ &= \frac{1}{W_n}\int x^i g(x)dx \\ &+\frac{1}{W_n}\sum_{s=1}^{n}\gamma_s^2\left(\int x^{2s+i}g(x)dx+\mu_s^2\int x^i g(x)dx-2\mu_s\int x^{s+i}g(x)dx\right) \\ &= \frac{1}{W_n}\left[\mu_i+\sum_{s=1}^{n}\gamma_s^2(\mu_{2s+i}+\mu_s^2\mu_i-2\mu_s\mu_{s+i})\right].\square \qquad \text{(A.5)} \end{aligned}$$

Proof of Proposition 7. The HPs, $H_s(x)$, can be rewritten in terms of the ME polynomials, $\Psi_s(x)$, as follows:

$$\begin{aligned} H_s(x) &= s!\sum_{k=0}^{[s/2]}\frac{(-1)^k}{k!(s-2k)!2^k}x^{s-2k} \\ &= s!\sum_{k=0}^{[s/2-1]}\frac{(-1)^k}{k!(s-2k)!2^k}\left[x^{s-2k}-\mu_{s-2k}^{+}\right]=s!\sum_{k=0}^{[s/2-1]}\frac{(-1)^k}{k!(s-2k)!2^k}\Psi_{s-2k}(x) \qquad \text{(A.6)} \end{aligned}$$

Then the GC can be expressed as a ME pdf as:

$$\begin{aligned} \pi(x,\mathbf{d}_n) &= \left(1+\sum_{s=1}^{n}d_sH_s(x)\right)\phi(x)=\left(1+\sum_{s=1}^{n}d_s s!\sum_{k=0}^{[s/2-1]}\frac{(-1)^k}{k!(s-2k)!2^k}\Psi_{s-2k}(x)\right)\phi(x) \\ &= \left(1+\sum_{s=1}^{n}\gamma_s\Psi_s(x)\right)\phi(x)=f_N(x,\boldsymbol{\gamma}_n).\square \qquad \text{(A.7)} \end{aligned}$$

Proof of Proposition 8. The GME cdf is given by

$$\begin{aligned} \int_{-\infty}^{a}f_N(x,\boldsymbol{\gamma}_n)dx &= \int_{-\infty}^{a}\left(1+\sum_{s=1}^{n}\gamma_s(x^s-\mu_s^{+})\right)\phi(x)dx \\ &= \int_{-\infty}^{a}\left(1-\sum_{s=1}^{n}\gamma_s\mu_s^{+}+\sum_{s=1}^{n}\gamma_s x^s\right)\phi(x)dx \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{a} \phi(x)dx \\
&- \sum_{s=1}^{n} \gamma_s \mu_s^+ \int_{-\infty}^{a} \phi(x)dx + \sum_{\substack{s=1 \\ s\ odd}}^{n} \gamma_s \int_{-\infty}^{a} x^s \phi(x)dx + \sum_{\substack{s=1 \\ s\ even}}^{n} \gamma_s \int_{-\infty}^{a} x^s \phi(x)dx \\
&= \int_{-\infty}^{a} \phi(x)dx \\
&- \sum_{s=1}^{n} \gamma_s \left[\left(x^{s-1} + (s-1)x^{s-3} + (s-1)(s-5)x^{s-5} + \ldots + \xi x^b \right) \phi(x)|_{-\infty}^{a} \right] \\
&= \int_{-\infty}^{a} \phi(x)dx - \sum_{s=1}^{n} \gamma_s \left(a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \xi a^b \right) \phi(a), \quad \text{(A.8)}
\end{aligned}
$$

where,

$$
\xi = \begin{cases} (s-1)(s-3)\cdots 2, & \forall s \text{ odd}, \\ (s-1)(s-3)\cdots 3, & \text{otherwise}, \end{cases} \qquad b = \begin{cases} 1, & \forall s \text{ even}, \\ 0, & \text{otherwise}. \end{cases} \quad \text{(A.9)}
$$

Note that the integrals are solved by parts as detailed below,

$$
\int x^s g(x)dx = \int x^{s-1} x g(x)dx = -x^{s-1} g(x) + (s-1) \int x^{s-2} \phi(x)dx \quad \text{(A.10)}
$$

since,

$$
\begin{aligned}
u = x^{s-1} &\Rightarrow du = (s-1)x^{s-2}dx, \\
dv = xg(x)dx &\Rightarrow v = \int x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} = -\phi(x). \quad \text{(A.11)}
\end{aligned}
$$

Therefore, by repeating the same argument recursively,

$$
\int x^s \phi(x)dx = \begin{cases} -\left[x^{s-1} + (s-1)x^{s-3} + (s-1)(s-3)x^{s-5} + \ldots + \xi \right] \phi(x), & \forall s \text{ odd} \\ \mu_s^+ \int \phi(x)dx \text{-} \left(x^{s-1} + (s-1)x^{s-3} + (s-1)(s-3)x^{s-5} + \ldots + \xi x \right) \phi(x), & \forall s \text{ even}, \end{cases} \quad \text{(A.12)}
$$

where $\mu_s^+ = \xi$. Furthermore, by applying recursively L'Hôpital's we obtain,

$$
\begin{aligned}
\lim_{x \to -\infty} [x^s \phi(x)] &= \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} \frac{x^s}{e^{\frac{1}{2}x^2}} = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} \frac{sx^{s-1}}{xe^{\frac{1}{2}x^2}} \\
&= \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} \frac{sx^{s-2}}{e^{\frac{1}{2}x^2}} = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} \frac{s(s-2)\cdots x}{xe^{\frac{1}{2}x^2}} \\
&= \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} \frac{s(s-2)\cdots 1}{e^{\frac{1}{2}x^2}} = 0. \square \quad \text{(A.13)}
\end{aligned}
$$

Proof of Proposition 9. The cdf of the GNT-GME in equation (9) can be obtained as follows

$$
\begin{aligned}
\int_{-\infty}^{a} F_N(x,\boldsymbol{\gamma}_n)dx &= \frac{1}{W_n}\int_{-\infty}^{a}\left(1+\sum_{s=1}^{n}\gamma_s^2(x^s-\mu_s^+)^2\right)\phi(x)dx \\
&= \frac{1}{W_n}\int_{-\infty}^{a}\left(1+\sum_{s=1}^{n}\gamma_s^2\mu_s^{+2}+\sum_{s=1}^{n}\gamma_s^2x^{2s}-2\sum_{s=1}^{n}\gamma_s^2\mu_s^+x^s\right)\phi(x)dx \\
&= \frac{1}{W_n}\int_{-\infty}^{a}\phi(x)dx+\frac{1}{W}\sum_{s=1}^{n}\gamma_s^2\mu_s^{+2}\int_{-\infty}^{a}\phi(x)dx+\frac{1}{W}\sum_{s=1}^{n}\gamma_s^2\int_{-\infty}^{a}x^{2s}\phi(x)dx \\
&\quad -2\frac{1}{W_n}\sum_{\substack{s=1\\ s\ odd}}^{n}\gamma_s\mu_s^+\int_{-\infty}^{a}x^s\phi(x)dx-2\frac{1}{W}\sum_{\substack{s=1\\ s\ even}}^{n}\gamma_s\mu_s^+\int_{-\infty}^{a}x^s\phi(x)dx \\
&= \frac{1}{W_n}\int_{-\infty}^{a}\phi(x)dx+\frac{1}{W}\sum_{s=1}^{n}\gamma_s^2\mu_s^{+2}\int_{-\infty}^{a}\phi(x)dx \\
&\quad -\frac{1}{W_n}\sum_{s=1}^{n}\gamma_s^2\left[(x^{2s-1}+(2s-1)x^{2s-3}+(2s-1)(2s-3)x^{2s-5}+\ldots+\mu_{2s}^+x)\phi(x)|_{-\infty}^{a}\right] \\
&\quad +\frac{2}{W_n}\sum_{s=1}^{n}\gamma_s^2\left[(x^{s-1}+(s-1)x^{s-3}+(s-1)(s-3)x^{s-5}+\ldots+\zeta x^b)\phi(x)|_{-\infty}^{a}\right] \\
&\quad +\frac{1}{W_n}\sum_{s=1}^{n}\gamma_s^2\mu_{2s}^+\int_{-\infty}^{a}\phi(x)dx-2\frac{1}{W}\sum_{s=1}^{n}\gamma_s^2\mu_s^{+2}\int_{-\infty}^{a}\phi(x)dx \\
&= \int_{-\infty}^{a}\phi(x)dx \\
&\quad +\frac{2}{W_n}\sum_{s=1}^{n}\gamma_s^2\mu_s^+\left[(x^{s-1}+(s-1)x^{s-3}+(s-1)(s-3)x^{s-5}+\ldots+\zeta x^b)\phi(x)|_{-\infty}^{a}\right] \\
&\quad -\frac{1}{W_n}\sum_{s=1}^{n}\gamma_s^2\left[(x^{2s-1}+(2s-1)x^{2s-3}+(2s-1)(2s-3)x^{2s-5}+\ldots+\mu_{2s}^+x)\phi(x)|_{-\infty}^{a}\right] \\
&= \int_{-\infty}^{a}\phi(x)dx \\
&\quad +2\frac{1}{W_n}\sum_{s=1}^{n}\gamma_s^2\mu_s^+(a^{s-1}+(s-1)a^{s-3}+(s-1)(s-3)a^{s-5}+\ldots+\zeta a^b)\phi(a) \\
&\quad -\frac{1}{W_n}\sum_{s=1}^{n}\gamma_s^2(a^{2s-1}+(2s-1)a^{2s-3}+(s-1)(s-3)a^{2s-5}+\ldots+\mu_{2s}^+a)\phi(a),
\end{aligned} \tag{A.14}
$$

where,

$$
\zeta=\begin{cases}(s-1)(s-3)\cdots 2, & \forall s \text{ odd},\\ (s-1)(s-3)\cdots 3a, & \text{otherwise},\end{cases}\qquad b=\begin{cases}1, & \forall s \text{ even},\\ 0, & \text{otherwise.}\square\end{cases} \tag{A.15}
$$

References

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

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

[3] Awartani, B. M. A. & Corradi, V. (2005). Predicting the volatility of the S&P-500 stock index via GARCH models: The role of asymmetries. *International Journal of Forecasting,* 21, 167-183.

[4] Bao, Y., Lee, T. H. & Saltoglu, B. (2006). Evaluating predictive performance of value-at-risk models in emerging markets: A reality check. *Journal of Forecasting*, 25, 101-128.

[5] Barton, D. E. & Dennis, K. E. R. (1952). The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal. *Biometrika,* 39, 425-427.

[6] Bekaert, G., Engstrom, E. & Ermolov, A. (2014). Bad environments, good environments: A non-Gaussian asymmetric volatility model. *Journal of Econometrics,* 186, 258-275.

[7] Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics,* 31, 307-327.

[8] Bollerslev, T. (1987) A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics,* 69, 542-547.

[9] Bollerslev, T. & Wooldridge, J. M. (1992). Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews,* 11, 143-172.

[10] Charlier, C. V. (1905). Uber die darstellung willkurlicher funktionen. *Arvik fur Mathematik Astronomi och fysik,* 9, 1-13.

[11] Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review,* 39, 841-862.

[12] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance,* 1, 223-236.

[13] Cramér, H. (1925). On some classes of series used in mathematical statistics. *Proceedings of the Sixth Scandinavian Congress of Mathematicians,* 399-425.

[14] Del Brio, E. B. Ñíguez, T. M. & Perote, J. (2009). Gram-Charlier densities: A multivariate approach. *Quantitative Finance,* 9, 855-868.

[15] Del Brio, E. B. Ñíguez, T. M. & Perote, J. (2011). Multivariate semi-nonparametric densities with dynamic conditional correlations. *International Journal of Forecasting,* 27, 347-364.

[16] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.

[17] Engle, R. F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives,* 15, 157-168.

[18] Engle, R. F. & Gonzalez-Rivera, G. (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics,* 9, 345-359.

[19] Engle, R. F. & Manganelli, S. (2004). CAViaR: Conditional value at risk by regression quantiles. *Journal of Business and Economic Statistics,* 22, 367-381.

[20] Ergen, I. (2014). Two-step methods in VaR prediction and the importance of fat tails. *Quantitative Finance,* 15, 1013-1030.

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

[22] Gallant, A. R. & Tauchen, G. (1989). Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications. *Econometrica*, 57, 1091-1120.

[23] Giacomini, R. & White, H. (2006). Test of conditional predictive ability. *Econometrica*, 74, 1545-1578.

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

[25] Göncü, A. & Yang. H. (2016). Variance-Gamma and Normal-Inverse Gaussian models: Goodness-of-fit to Chinese high-frequency index returns. *The North American Journal of Economics and Finance*, 36, 279-292.

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

[27] Huang, H. H., Lin, S. H., Wang, C. P. & Chiu, C. Y. (2014). Adjusting MV-efficient portfolio frontier bias for skewed and non-mesokurtic returns. *The North American Journal of Economics and Finance*, 29, 59-83

[28] Jensen, M. B. & Lunde, A. (2001). The NIG-S&ARCH model: A fat-tailed, stochastic and autoregressive conditional heteroskedastic volatility model. *Econometrics Journal,* 4, 319-342.

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

[30] Kuester, K., Mittnik, S. & Paolella, M. S. (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics*, 4, 53-89.

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

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

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

[34] Lin, S. H., Huang, H. H. & Li, S. H. (2015). Option pricing under truncated Gram–Charlier expansion. *The North American Journal of Economics and Finance*, 32, Pages 77-97.

[35] López, J. A. (1999). Methods for evaluating value-at-risk estimates. *FRBSF Economic Review,* 2, 3-17.

[36] Maheu, J. M. & McCurdy, T. H. (2011). Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160, 69-76.

[37] Mauleón, I. & Perote, J. (2000). Testing densities with financial data: An empirical comparison of the Edgeworth-Sargan density to the Student's t. *European Journal of Finance,* 6, 225-239.

[38] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica,* 59, 347-370.

[39] Ñíguez, T. M. & Perote, J. (2016). Multivariate moments expansion density: Application of the dynamic equicorrelation model. *Journal of Banking and Finance*, 72, S216–S232.

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

[41] Patton, A. (2010). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160, 246-256.

[42] Sargan, J. D. (1976). Econometric estimators and the Edgeworth approximation. *Econometrica,* 44, 421-448.

[43] Terasvirta, T. (2009). *An introduction to univariate GARCH models.* In Handbook of Financial Time Series, T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch, (eds.). New York: Springer.

Tables

TABLE 1

Estimation results

	Normal	GC	Sk-GC	GME	Student's t	NIG
			Panel 1: S&P 500			
φ_0	0.0109 (1.82)	0.0102 (1.37)	0.0110 (1.42)	0.0066 (1.26)	0.0106 (1.67)	0.0103 (1.71)
φ_1	0.0381 (4.24)	0.0491 (3.72)	0.0521 (3.81)	0.0400 (4.35)	0.0381 (3.79)	0.0379 (3.88)
φ_2	0.9526 (75.7)	0.9554 (80.4)	0.9545 (78.6)	0.9554 (80.5)	0.9536 (67.8)	0.9537 (70.1)
ξ	0.1839 (2.20)	0.1109 (1.54)	0.1176 (1.69)	0.1331 (1.70)	0.1119 (1.57)	0.1148 (1.61)
γ_2		-0.0641 (-2.05)	-0.0731 (-2.56)	0.0388 (0.25)		
γ_3			-0.0146 (-3.53)			
γ_4		0.0182 (6.60)	0.0177 (-11.45)	0.0157 (8.23)		
ν					6.1789 (8.53)	
ζ						1.6791 (5.71)
AIC	2.9653	2.9245	2.9209	2.9372	2.9139	2.9137
			Panel 2: FX £/$			
φ_0	0.0034 (2.22)	0.0030 (2.31)	0.0029 (2.32)	0.0041 (2.27)	0.0024 (1.69)	0.0026 (1.84)
φ_1	0.0390 (5.10)	0.0364 (5.35)	0.0358 (5.30)	0.0437 (5.15)	0.0490 (5.32)	0.0487 (5.40)
φ_2	0.9498 (82.8)	0.9430 (77.5)	0.9437 (78.1)	0.9431 (72.7)	0.9474 (84.8)	0.9451 (80.8)
ξ	0.0178 (0.25)	0.0267 (0.38)	0.0264 (0.38)	0.0312 (0.44)	0.0143 (0.23)	0.0194 (0.32)
γ_2		0.0267 (1.54)	0.0267 (1.56)	0.1260 (3.85)		
γ_3			-0.0034 (-1.06)			
γ_4		0.0238 (12.6)	0.0237 (-12.6)	0.0225 (13.2)		
ζ						1.0790 (8.43)
ν					4.6813 (13.3)	
AIC	1.5936	1.5353	1.5355	1.5440	1.5320	1.5266

TABLE 1 *(continued)*

	Normal	GC	Sk-GC	GME	Student's t	NIG
			Panel 3: Microsoft			
φ_0	0.0728 (3.15)	0.0710 (2.70)	0.0706 (2.73)	0.0801 (2.27)	0.0405 (2.92)	0.0423 (2.96)
φ_1	0.1000 (4.86)	0.1290 (4.15)	0.1287 (4.21)	0.1173 (5.15)	0.1081 (5.31)	0.1065 (5.30)
φ_2	0.8772 (40.1)	0.8734 (35.5)	0.8747 (36.7)	0.8613 (72.7)	0.8860 (47.8)	0.8849 (46.8)
ξ	0.2489 (3.53)	0.2234 (4.38)	0.2225 (4.08)	0.2071 (2.88)	0.2811 (5.69)	0.2754 (5.43)
γ_2		-0.0592 (-2.27)	-0.0607 (-2.34)	0.0413 (3.85)		
γ_3			-0.0037 (-1.10)			
γ_4		0.0205 (9.86)	0.0205 (9.81)	0.0181 (11.3)		
ζ						1.2522 (7.13)
ν					5.1164 (11.3)	
AIC	3.8418	3.7875	3.7877	3.8044	3.7711	3.7875

Notes: This table presents the estimation results (BW (Q)ML t-statistics in brackets) for the Gaussian, Student's t, GC, Sk-GC, GME and NIG AR(1)-AGARCH(1,1) models. AIC denotes the Akaike Information Criterion statistic. γ_s accounts for the weighting parameter of the $s-th$ order Hermite polynomial in the SNP distributions. ν captures the degrees of freedom parameter in the Student's t distribution, and ζ is the steepness parameter in the NIG distribution.

TABLE 2

Out-of-sample volatility forecasting performance

	Normal	GC	Sk-GC	GME	Student's t
Panel 1: S&P500					
GC	1.56				
Sk-GC	0.02	-1.78*			
GME	0.23	-1.82*	0.46		
Student's t	-0.23	-1.90*	-0.65	-0.77	
NIG	-0.19	-1.85*	-0.522	-0.63	0.45
Panel 2: FX £/$					
GC	1.98**				
Sk-GC	0.38	-1.38			
GME	0.90	-1.48	0.503		
Student's t	-0.28	-1.06	-0.74	-0.70	
NIG	0.09	-1.05	-0.58	-0.58	0.99
Panel 3: Microsoft					
GC	-1.68*				
Sk-GC	-0.27	-1.81*			
GME	0.18	-1.70*	0.65		
Student's t	-0.06	-2.16**	0.27	-0.12	
NIG	-0.19	-2.12**	-0.11	-0.27	-0.74

Notes: This table presents the GW test t-statistics for the difference in the MSFE of the model in the row versus the model in the column. A negative statistic indicates that the model in column presents a higher MSFE. * denotes rejection of the null hypothesis of no difference in MSFE at 10% level, ** indicates rejection at 5% level. Predictions 1,000.

TABLE 3

Models evaluation for VaR predictive accuracy

	α	$\widehat{\alpha}$	M_α	HIT^{p-v}	LR$_{uc}^{p-v}$	LR$_{cc}^{p-v}$
			Panel 1: S&P500			
Normal	.01	.018	28.7	.04**	.02**	.04*
	.05	.057	96.5	.14	.32	.55
	.1	.115	190	.09	.29	.21
Normal^{E}	.01	.010	13.1	.99	.99	.90
	.05	.057	98.6	.13	.32	.55
	.1	.120	209	.06*	.04*	.06*
GC	.01	.017	26.8	.13	.04	.09
	.05	.069	119	.02	.00	.03
	.1	.131	238	.00	.00**	.00**
GC-Sk	.01	.018	27,1	.04	.02**	.06*
	.05	.074	129	.00	.00	.00
	.1	.146	263	.00	.00	.00
GME	.01	.016	20.8	.13	.07*	.16
	.05	.060	81.7	.23	.15	.16
	.1	.125	170	.07*	.01**	.01**
Student's t	.01	.003	4.20	.88	.58	.11
	.05	.034	59.1	.50	.02	.04
	.1	,082	144	.58	.01	.03
NIG	.01	.011	17.0	.99	.75	.84
	.05	.060	102	.11**	.15*	.36**
	.1	.126	122	.03**	.00**	.02**

TABLE 3 (*continued*)

	α	$\widehat{\alpha}$	M_α	HIT^{p-v}	LR_{uc}^{p-v}	LR_{cc}^{p-v}
Panel 2: FX £/$						
Normal	.01	.018	25.7	.02	.02**	.05*
	.05	.060	81.9	.22	.15	.16
	.1	.115	155	.31	.12	.10
Normal^{E}	.01	.014	18.4	.12	.23	.39
	.05	.060	81.8	.22	.15	.16
	.1	.126	170	.07*	.00**	.01*
GC	.01	.003	4.47	.89	.00	.03
	.05	.024	34.5	.08	.00	.11
	.1	.082	109	.29	.05	.01
GC-Sk	.01	.079	4.40	.89	.00**	.03**
	.05	.024	34	.08*	.00**	.00**
	.1	.003	105	.18	.02**	.01*
GME	.01	.011	16.9	.99	.75	.84
	.05	.056	96.8	.19	.39	.61
	.1	.120	208	.09*	.04**	.09*
Student's t	.01	.003	4.36	.12	.23	.39
	.05	.027	38	.22	.15	.16
	.1	.079	106	.07	.00	.01
NIG	.01	.010	17.8	.55	.23	.39
	.05	.064	86.1	.03	.05	.02
	.1	.133	180	.08	.00**	.00**

TABLE 3 (*continued*)

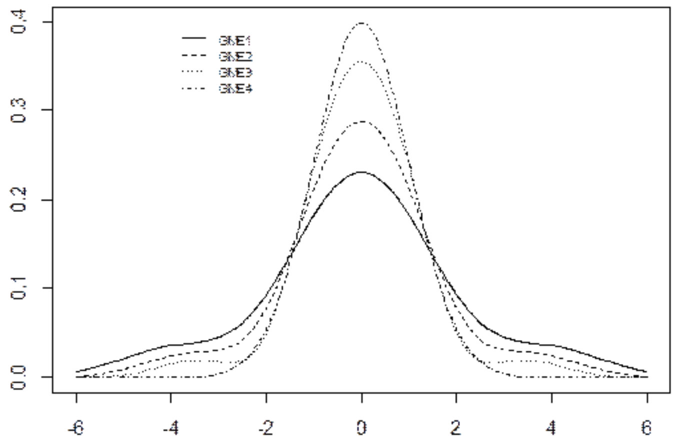
	α	$\widehat{\alpha}$	M_α	HIT^{p-v}	LR_{uc}^{p-v}	LR_{cc}^{p-v}
			Panel 3: Microsoft			
Normal	.01	.014	29.9	.13	.23	.39
	.05	.049	97.7	.47	.88	.95
	.1	.088	179	.76	.19	.41
Normal^E	.01	.010	18.4	.40	.90	.99
	.05	.054	105	.43	.70	.56
	.1	.102	210	.21	.96	.83
GC	.01	.019	38.4	.03*	.01	.02
	.05	.077	157	.00	.00	.00
	.1	.136	282	.00	.00	.00
GC-Sk	.01	.015	31.5	.14**	.26**	.13**
	.05	.074	149	.00**	.00**	.01**
	.1	.134	275	.00	.00	.00
GME	.01	.013	23.7	.11	.36	.55
	.05	.055	108	.25	.47	.62
	.1	.099	202	.32	.91	.95
Student's t	.01	.003	7.95	.89	.00	.03
	.05	.027	62.6	.13	.00	.00
	.1	.068	142	.02	.00	.00
NIG	.01	.013	22.5	.11	.55	.36
	.05	.056	107	.55	.53	.39
	.1	.105	214	.20	.81	.60

Notes: This table gathers the results of statistical tests and loss function values for one-step-ahead VaR forecast from Normal, Normal^E, Student's t, GC-S, Sk-GC, GME and NIG AGARCH models. $\widehat{\alpha}$ denotes the estimated unconditional coverage probability. M_α stands for the magnitude of the exception statistic. HIT^{p-v} is the p-value corresponding to the Engle and Manganelli (2004) test. LR_{uc}^{p-v} and LR_{cc}^{p-v} denote the p-value of the unconditional and conditional coverage LR tests (Christoffersen, 1998). The asterisk represents rejection of the null hypothesis at least at 5% level. Predictions 1,000.

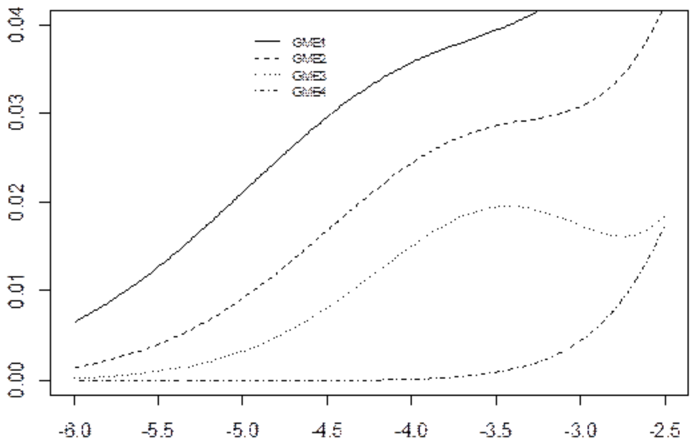
Figures

FIGURE 1
ME density shapes

Panel A: Full density

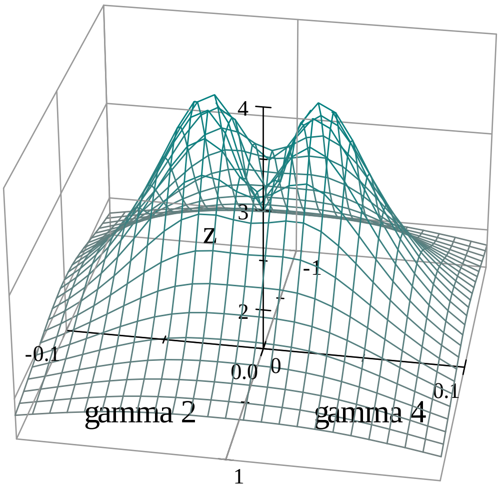


Panel B: Left tail

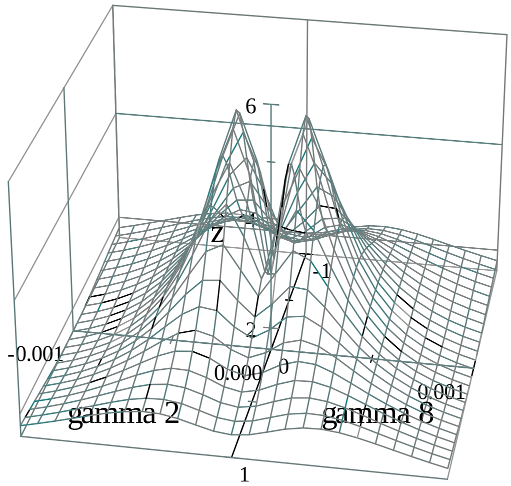


Notes: This Figure illustrates the allowable shapes of GME densities. GME1 denotes GME for $\gamma_4 = -0.01$, $\gamma_6 = 0.005$, $\gamma_8 = 0.001$; GME2 denotes GME for $\gamma_4 = -0.05$, $\gamma_6 = 0.01$, $\gamma_8 = 0.0005$; GME3 denotes GME for $\gamma_4 = -0.08$, $\gamma_6 = 0.01$, $\gamma_8 = 0.008$; GME4 denotes GME $\gamma_4 = 0$, $\gamma_6 = 0$, $\gamma_8 = 0$. All pictured GME densities are symmetric ($\gamma_s = 0\ \forall s$ odd) and truncated at $n = 8$.

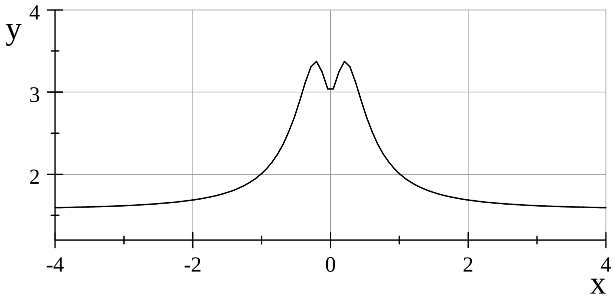
FIGURE 2
Range of kurtosis for symmetric GNT GME densities



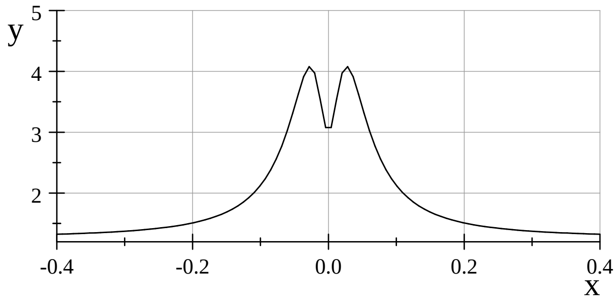
GME n=4, Range=[1.258,4.082]



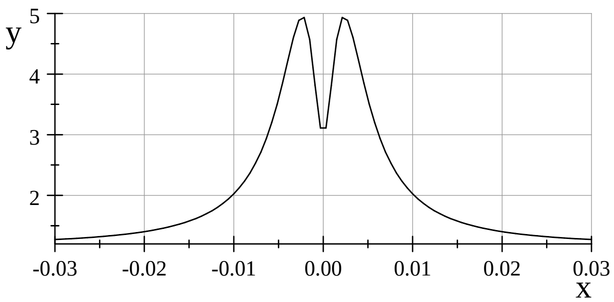
GME n=8 , γ_4=γ_6=0, Range=[1.2,5.585]



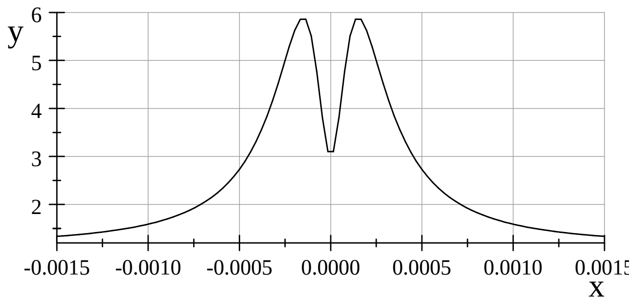
GME n=2 , Range=[1.56,3.375]



GME n=4 , γ_2=0, Range=[1.258,4.082]



GME n=6, γ_2=γ_4=0, Range=[1.163,4.954]

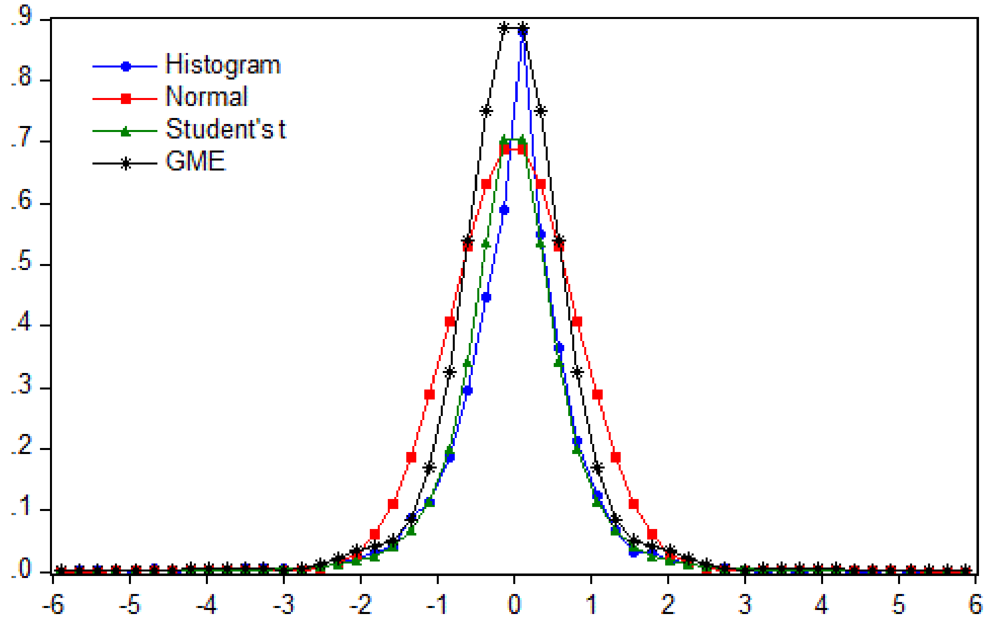


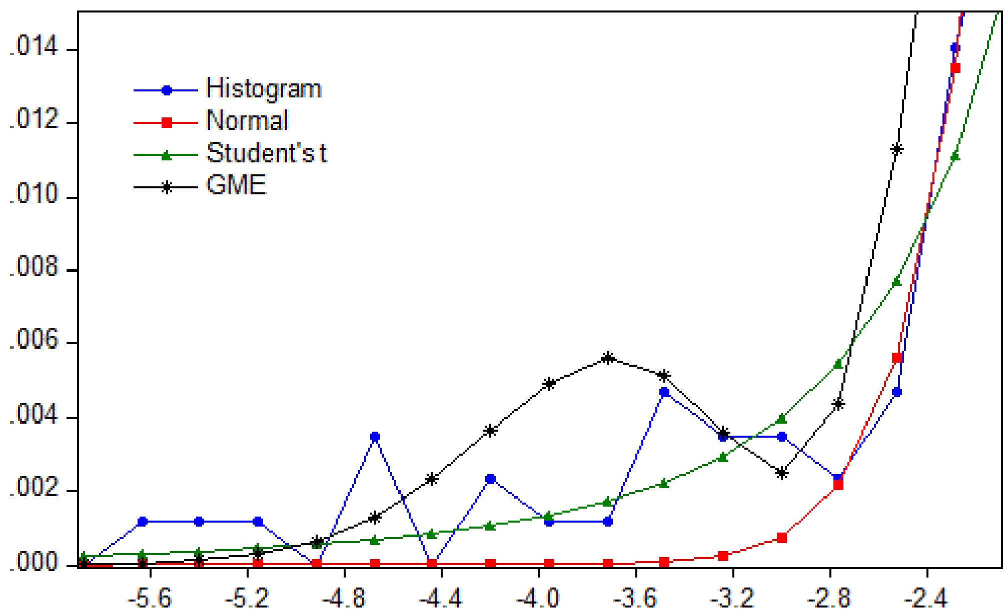
GME $n = 8$, γ_2=γ_4=γ_6=0, Range=[1.12,5.892]

Notes: This Figure provides the kurtosis range allowed by symmetric (γ_s= 0 $\forall s$ odd) GNT GME densities.

FIGURE 3

Empirical and estimated distributions

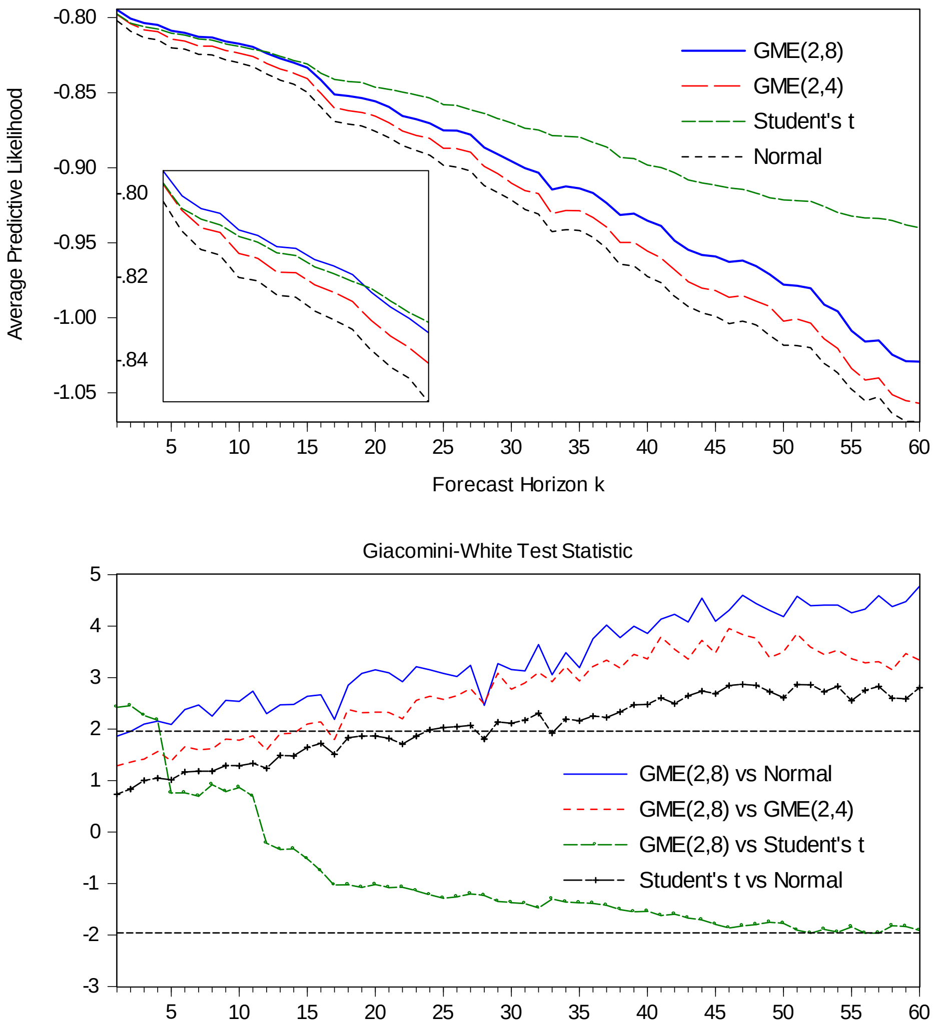




Notes: This Figure plots the empirical and estimated Normal, Student's t and GNT-GME distributions for the FX £/$ return series.

FIGURE 4

Multiperiod density forecast



Notes: This Figure plots average predictive likelihood and Giacomini-White (GW) test statistic for daily FX \$/£ returns as functions of forecast horizon $k = 1, 2, ..., 60$ days. Models with better forecasting performance for every k provide higher average predictive likelihood. The GW test statistic measures the significance of the differences between the predictive likelihood of two alternative models. A positive statistic for "i vs j" is evidence in favor of model i against model j. GME denotes symmetric GNT GNE; (2,8) denotes $n = 8$, $\gamma_4 = \gamma_6 = 0$, and (2,4) denotes $n = 4$. Dashed horizontal lines correspond to 5% significance levels.