The General Moments Expansion: an application for financial risk*

Trino-Manuel Ñíguez† Javier Perote‡
University of Westminster Rey Juan Carlos University

Abstract

This paper presents a family of semi-nonparametric distributions based on what we name General Moments Expansions (GME). We analyze their theoretical properties and argue that GME densities are simpler and more general than the well-known Edgeworth and Gram-Charlier (EGC) densities since their application is straightforward; it only requires that the expanded density has finite moments up to the truncation order. We show that GME densities preserve the flexibility proper of densities based on EGC expansions for modelling asset returns. The performance of a GME of the Normal density is compared to both the Gaussian and the Student’s t by means of an empirical application for forecasting financial risk.

Keywords: Density expansions; Edgeworth and Gram-Charlier distributions; Financial risk forecasting.

JEL classification: C16, C53, G12.

MS classification: 62E17, 33C45.

*This research has been supported by the Spanish Ministry of Education under grant SEJ2006-06104/ECON. We wish to thank Antonio Rubia and participants at the 27th International Symposium on Forecasting in New York, USA, June 2007, XXXII Symposium of Economic Analysis in Granada, Spain, December 2007, and 2nd International Workshop on Computational and Financial Econometrics in Neuchâtel, Switzerland, June 2008, for their comments. All remaining errors are ours.

†Department of Economics and Quantitative Methods, Westminster Business School, University of Westminster, 35 Marylebone Road, NW1 5LS London, UK. Phone: +44 (0) 22 7911 5000. Fax: +44 (0) 22 7911 5839. E-mail: T.M.Niguez@wmin.ac.uk

‡Corresponding author: Department of Economía Aplicada II y Fundamentos del Análisis Económico, Facultad de Ciencias Jurídicas y Sociales, Universidad Rey Juan Carlos, Campus de Vicálvaro, Pº de los Artilleros s/n, 28032 Madrid, Spain. Phone: +34 91 488 7684. Fax: +34 91 488 7145. E-mail: javier.perote@urjc.es
1 Introduction

During the last decades, the financial econometrics literature related to the study of heavy-tailed distributions has undergone a huge development. Among the different used methodologies, those of the most fruitful have been the semi-nonparametric techniques brought into econometrics by Sargan (1975, 1976) and later developed by authors such as: Jarrow and Rudd (1982), Gallant and Tauchen (1989), Corrado and Su (1997), Harvey and Siddique (1999), Mauleón and Perote (2000), Nishiyama and Robinson (2000), Jondeau and Rockinger (2001), Velasco and Robinson (2001), Ñíguez and Perote (2004), Verhoeven and McAleer (2004), and León et al. (2004, 2005), among others. Those papers use the property of the infinite Edgeworth and Gram-Charlier series to define semi-nonparametric densities, which are characterized by their flexibility admitting as much parameters as necessary to parsimoniously approximate a true target distribution. The resulting models have been used for many purposes, from hypotheses testing to density and value-at-risk (VaR hereafter) forecasting.

Truncated Edgeworth and Gram-Charlier (EGC hereafter) series expansions of the Normal density, so-called EGC densities, can be shown as well-defined densities because of the orthogonality property of the Hermite polynomials. In particular, that property is the basis for proving that: i) EGC densities integrate to one, ii) the EGC density parameters can be expressed in terms of its moments (or cumulants) and, iii) the cumulative distribution function (cdf hereafter) has a closed analytical expression.

Nevertheless, the EGC densities may present computational problems — e.g. when estimating large expansions, introducing time-varying structures for other conditional moments than the variance or generalising the distribution to a multivariate framework. Hence, they are considered more complex to implement empirically than parametric distributions. In this paper we address these issues by proposing a different type of polynomial expansions, that we name General Moments Expansions (GME hereafter) to define semi-nonparametric densities that preserve both the flexibility and good performance of the EGC distributions but are easier to analyze and implement and more general, because i) they embody simpler than Hermite (orthogonal) polynomials, what simplifies the model
specification, and \( ii \) they can be very straightforwardly applied to expand any density, only requiring the condition of finite moments up to the chosen truncation order.

The goodness-of-fit of a GME of the Normal density is tested by means of an empirical application for forecasting the conditional variance and VaR of exchange-rates returns, in comparison with Student’s t and Gaussian distributions. The out-of-sample forecasts are evaluated by using the following criteria: \( i \) the ranking-robust loss functions for imperfect volatility proxies analyzed in Patton (2006), \( ii \) the VaR predictive accuracy criteria in López (1999), and \( iii \) the predictive quantile loss function (Koenker and Bassett, 1978). The results obtained from this application are twofold: First, GME and Gaussian models provide statistically similar performance for forecasting the conditional variance but both are preferred to the Student’s t model. Second, the GME model yields more accurate VaR forecasts than the Gaussian VaR model of Engle (2001) and the Student’s t.

The remainder of the paper is organized as follows. Section 2 reviews the densities based on Hermite polynomial expansions and introduces the GME analysing their theoretical properties. In Section 3 the performance of the GME densities is tested through an empirical application for forecasting financial risk and, Section 4 summarizes the conclusions.

\section{Density Expansions}

\subsection{The Edgeworth and Gram-Charlier distribution}

Let \( x \) be a random variable whose density is defined in terms of the Type A series expansions introduced in the seminal papers of Edgeworth (1896, 1907) and Charlier (1905). Within this family of densities is the so-called EGC pdf (equation (1)),

\[ \pi(x, d) = \left( 1 + \sum_{s=1}^{n} d_s H_s(x) \right) \phi(x), \]  

where \( \phi(\cdot) \) denotes the standard Normal pdf, \( d = (d_1, d_2, \ldots, d_n)' \in \mathbb{R}^n \), and \( H_s(\cdot) \) is the Hermite polynomial of order \( s \), which can be defined in terms of the derivatives of \( \phi(\cdot) \) as,

\[ \frac{d^s \phi(x)}{dx^s} = (-1)^s H_s(x) \phi(x). \]  

3
For the sake of the analysis that follows, it is worth recapitulating that Hermite polynomials satisfy the orthogonality properties in equations (3) and (4); see Kendall and Stuart (1977) for further details,

\[
\int H_s(x) H_j(x) \phi(x) dx = 0, \quad \forall s, j \geq 0, \quad s \neq j, \tag{3}
\]

\[
\int H_s(x)^2 \phi(x) dx = s!, \quad \forall s \geq 0. \tag{4}
\]

On the other hand, note that the density in equation (1) is based on a EGC series truncated at order \( n \) and, consequently, the function may take negative values for some values of its parameters. So for \( \pi(x, d) \) being well defined, positivity must be guaranteed, this can be done in different ways: \( i \) Mauleón and Perote (2000) proposed controlled optimization focusing on appropriate election of starting values, \( ii \) Jondeau and Rockinger (2001) proposed parameter constraints, and \( iii \) León et al. (2004, 2005) and Ñíguez and Perote (2004) focused on reformulations based on the methodology in Gallant and Nychka (1989) and Gallant and Tauchen (1989). In particular, the latter proposed the following reformulation, named Positive Edgeworth-Sargan (PES),

\[
\Pi(x, d) = \frac{1}{\hat{w}} \left( 1 + \sum_{s=1}^{n} d^2_s H_s(x)^2 \right) \phi(x), \tag{5}
\]

and the former proposed squaring the sum of polynomials,

\[
\bar{\Pi}(x, d) = \frac{1}{\bar{w}} \left( 1 + \sum_{s=1}^{n} d_s H_s(x) \right)^2 \phi(x), \tag{6}
\]

where \( w, \hat{w} \) are the constants that guarantee that the densities integrates to one,

\[
w = \int \left( 1 + \sum_{s=1}^{n} d^2_s H_s(x)^2 \right) \phi(x) dx = 1 + \sum_{s=1}^{n} d^2_s s!, \tag{7}
\]

\[
\hat{w} = \int \left( 1 + \sum_{s=1}^{n} d_s H_s(x) \right)^2 \phi(x) dx = 1 + \sum_{s=1}^{n} \frac{(d_s - \mu^+_s)^2}{s!}, \tag{8}
\]

where \( \mu^+_s \) is the \( s \)-th moment of \( \phi(x) \).
2.1.1 Notation

Let \( H = [H_1(x), H_2(x), \ldots, H_n(x)]' \in \mathbb{R}^n \), then the density described in equation (1) can be alternatively defined in matrix form as,

\[
\pi(x, d) = (1 + H'd) \phi(x),
\]

and the orthogonality conditions, in equations (3) and (4), can be summarized in the diagonal matrix \( S \) as,

\[
S = \int HH'\phi(x)dx = \text{diag} \{1!, 2!, \ldots, n!\}.
\]

Furthermore, a closed-form for the vector of Hermite polynomials can be obtained as,

\[
H = BZ + I'\mu^+,
\]

where \( B \) is a triangular matrix of order \( n \), containing the coefficients for the different powers of \( x \) in \( H \), which, in turns, form the vector \( Z \in \mathbb{R}^n \), \( \mu^+ \in \mathbb{R}^n \) accounts for the first \( n - th \) moments of \( \phi(\cdot) \), and \( I^* \) is a diagonal matrix of order \( n \) that contains the sign of the corresponding intercept of every Hermite polynomial. Note that, without loss of generality, the matrices \( B \) and \( I^* \), and the vector \( \mu^+ \) in equation (12) below, are written for \( n \) even.

\[
B = \begin{bmatrix}
1 & & & & 0 \\
0 & 1 & & & \\
-3 & 0 & \ddots & & \\
0 & -6 & 0 & \ddots & \\
\vdots & \vdots & \vdots & \ddots & \ddots \\
0 & 0 & \frac{(-1)^{n-2}n!}{2^{n-2} n-2!2!} & \frac{(-1)^{n-4}n!}{2^{n-4} n-4!4!} & \cdots & 0 & 1
\end{bmatrix},
\]

\[
Z = \begin{bmatrix}
x \\
x^2 \\
\vdots \\
x^n
\end{bmatrix}, \quad \mu^+ = \begin{bmatrix}
\mu_1^+ \\
\mu_2^+ \\
\vdots \\
\mu_n^+
\end{bmatrix}, \quad I^* = \begin{bmatrix}
0 & 0 \\
-1 & 0 \\
0 & 1 \\
\vdots & \ddots \\
0 & \frac{(-1)^{n/2}}{2^{n/2} n^{n/2}!}
\end{bmatrix}.
\]
2.2 The General Moments Expansions

The Edgeworth and Gram-Charlier expansions can theoretically be applied to expand not only the Normal density, but also any continuous and differentiable pdf. The resulting densities admit generalized autoregressive conditional heteroscedasticity (GARCH) type of specifications (Engle (1982) and Bollerslev (1986)) to account for possible clustering dynamics in conditional moments rather than the second one, e.g., conditional skewness and kurtosis; see Harvey and Siddique (1999) and León et al. (2005) for applications of EGC expansions to the Normal density in that context. However, in practice, guaranteeing positivity in those cases is not straightforward, specially when expanding other distributions rather than the Normal, due to the complexity of the resulting Hermite polynomials; as conjectured in Mauleón and Perote (2000). In this paper we propose a general class of expansions that could be used to define easily tractable densities in the aforementioned context. Such an expansions, that we name GME, are defined in terms of the moments of the distribution taken as basis, and are valid to apply to any distribution with finite moments up to the expansion truncation order.

Definition 1 The GME pdf using a chosen pdf, \( g(\cdot) \), as basis, with finite non-central moments \( E[x^s] = \mu_s \ \forall s = 1, 2, \ldots, n \), is defined as,

\[
f(x, \gamma) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s(x)\right) g(x),
\]

where \( \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)' \in \mathbb{R}^n \) is a vector of parameters, and \( \{\Psi_s(x)\}_{s=1}^{n} \) is a polynomials sequence of the form,

\[
\Psi_s(x) = x^s - \mu_s.
\]

The density \( f(x, \gamma) \) can be expressed in matrix form as,

\[
f(x, \gamma) = (1 + (Z - \mu)' \gamma) g(x).
\]

where \( Z \) is defined in equation (12), and \( \mu \in \mathbb{R}^n \) is a vector that contains the moments of \( g(\cdot) \).

The GME family of densities satisfies the following properties:

1. Up to one integration.
2. Sufficient conditions exist to guarantee positivity.

3. The GME density moments can be easily computed. As a consequence, it can be expressed in terms of its non-central moments.

4. The EGC density can be obtained from a suitable GME density as a particular case.

5. The standardized GME density can be computed by a simple linear transformation.

6. The GME pdf admits Gallant and Nychka’s (1987) type of transformations to ensure positivity.

7. The moments of the "Positive" GME density are linear functions of the squared density parameters.

8. The GME cdf can be easily obtained.

9. The "Positive" GME cdf can be straightforwardly worked out.

These properties are developed below, all proofs are provided in the appendix.

2.2.1 GME properties

**Proposition 1** \( f(x, \gamma) \) integrates to one: \( \int f(x, \gamma)dx = 1 \).

**Proposition 2** If \( 0 \leq \gamma_s \leq \frac{1}{n \mu_s} \) \( \forall s \) even, and \( \gamma_s = 0 \) \( \forall s \) odd, then \( f(x, \gamma) \) is positive.

**Proposition 3** The non-central moments of \( f(x, \gamma) \) can be easily computed from the moments of \( g(x) \) as,

\[
m_i = E[x^i] = \mu_i + \sum_{s=1}^{n} \gamma_s (\mu_{s+i} - \mu_s \mu_i), \quad \forall i = 1, 2, 3, ...
\]

Alternatively, the first \( n \)th non-central moments of \( f(x, \gamma) \) — in vector \( \mathbf{M} = (m_1, m_2, \ldots, m_n)' \), can be expressed in matrix form as,

\[
\mathbf{M} = \mathbf{\mu} + \mathbf{A} \gamma.
\]
where A is a symmetric matrix of order n whose $i$–th element is $\{a_{ij}\} = \{\mu_{s+i} - \mu_{s}\mu_{i}\}$.

**Corollary 1** If $g(x)$ is a symmetric density, the even and odd non-central moments of $f(x, \gamma)$ depend on its even and odd parameters, respectively. Therefore, the matrix A can be rewritten as a block diagonal matrix, where the submatrices $A_I$ and $A_{II}$ contain the odd and even parameters of $f(x, \gamma)$, respectively. Accordingly, the vectors $M$, $\mu$ and $\gamma$ can also be partitioned in $M_I$ and $M_{II}$, $\mu_I$ and $\mu_{II}$, and $\gamma_I$ and $\gamma_{II}$, respectively, containing the even and the odd moments of $f(x, \gamma)$, the density moments of $g(x)$, and the parameters of $f(x, \gamma)$, respectively. Equation (18) expresses equation (17) in terms of the resulting partitioned non-homogeneous equations system.

$$
\begin{bmatrix}
M_I \\
M_{II}
\end{bmatrix} =
\begin{bmatrix}
\mu_I \\
\mu_{II}
\end{bmatrix} +
\begin{bmatrix}
A_I & 0 \\
0 & A_{II}
\end{bmatrix}
\begin{bmatrix}
\gamma_I \\
\gamma_{II}
\end{bmatrix}.
$$

(18)

Furthermore, if A is full-rank, then the system in equation (17) or equation (18) has a trivial solution given by,

$$
\gamma = A^{-1}(M - \mu),
$$

(19)

$$
\gamma_I = A_I^{-1}(M_I - \mu_I) \quad \text{and} \quad \gamma_{II} = A_{II}^{-1}(M_{II} - \mu_{II}).
$$

(20)

**Corollary 2** $f(x, \gamma)$ can be expressed in terms of its first $n$–th non-central moments as,

$$
f(x, M) = [1 + (Z - \mu)'A^{-1}(M - \mu)] g(x) = \left(1 + \sum_{s=1}^{n}(m_s - \mu_s)\theta_s(x)\right) g(x),
$$

(21)

where $\theta_s(x)$ is the polynomial corresponding to the $s$–th element of the vector $(Z - \mu)'A^{-1}$.

To compare the GME density, $f(x, \gamma)$, and the EGC density, $\pi(x, d)$, Example 1 provides a GME of the standard Normal.

**Example 1:** Let $f_N(x, \gamma)$ be the GME pdf using the Normal density, $\phi(\cdot)$, as basis,

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

(22)

where,

$$
\mu_s^+ = \begin{cases} 
\frac{s!}{2^s(s/2)!} = (s - 1)(s - 3)(s - 4) \cdots 3 & \forall s \text{ even} \\
0 & \text{otherwise}
\end{cases}
$$

(23)
The density \( f_N(x, \gamma) \) is always positive and its moments can easily be computed as shown in equation (24), providing that: \( 0 \leq \gamma_s \leq \frac{2^s(s/2)!}{s!} \) \( \forall s \) even, and \( \gamma_s = 0 \) \( \forall s \) odd.

\[
E[x'] = \mu_i^+ + \sum_{s=1}^{n} \gamma_s (\mu_{s+i}^+ - \mu_s^+ \mu_i^+) = \frac{i!}{2^s} \left[ 1 + \sum_{s=1}^{n} \gamma_s \frac{s!}{i!} \frac{1}{2^s} \left( \frac{(s+i)!}{(s+s)!} - \frac{i!}{2^s} \right) \right].
\] (24)

Specifically, if we consider the \( f_N(x, \gamma) \) pdf constrained to \( \gamma_s = 0 \) \( \forall s \geq 5 \) — as in equation (25) below, the first four moments of the resulting truncated density, denoted as \( \tilde{f}_N(x, \gamma) \), are given in equation (26).

\[
\tilde{f}_N(x, \gamma) = \left[ 1 + \gamma_1 x + \gamma_2 (x^2 - 1) + \gamma_3 x^3 + \gamma_4 (x^4 - 3) \right] \phi(x).
\] (25)

\[
E[x] = m_1 = \gamma_1 + 3 \gamma_3,
E[x^2] = m_2 = 1 + 2 \gamma_2 + 12 \gamma_4,
E[x^3] = m_3 = 3 \gamma_1 + 15 \gamma_3,
E[x^4] = m_4 = 3 + 12 \gamma_2 + 96 \gamma_4.
\] (26)

This distribution can be expressed in terms of its first four moments as,

\[
\tilde{f}_N(x, \mathbf{M}) = [1 + m_1 \theta_1(x) + (m_2 - 1) \theta_2(x) + m_3 \theta_3(x) + (m_4 - 3) \theta_4(x)] \phi(x),
\] (27)

where,

\[
\theta_1(x) = \frac{x(5 - x^2)}{2},
\theta_2(x) = 2(x^2 - 1) - \frac{x^4 - 3}{4},
\theta_3(x) = \frac{x(x^2 + 3)}{6},
\theta_4(x) = \frac{x^4 - 3}{24} - \frac{x^2 - 1}{4}.
\] (28)

**Proposition 4** If \( \gamma = A^{-1}B^{-1}(S \times \mathbf{d}) - A^{-1}(B^{-1} + \Gamma')\mu^+ \), the Normal GME density, \( f_N(x, \gamma) \), and the EGC density, \( \pi(x, \mathbf{d}) \), have the same moments.

**Proposition 5** If \( m_i^* = E[(x - m_1)^i] \ \forall i = 1, 2, \ldots \), are the central moments of \( f(x, \gamma) \), then the standardized GME density — i.e. zero mean and variance one, denoted as \( f(z, \cdot) \),
can be defined either in terms of the density parameters, $f(z, \gamma)$, or in terms of the density moments, $f(z, M)$.

$$f(z, \gamma) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s \left(m_2^{s1/2} z + m_1\right)\right) g \left(m_2^{s1/2} z + m_1\right) m_2^{s1/2}, \quad (29)$$

$$f(z, M) = \left(1 + \sum_{s=1}^{n} (m_s - \mu_s) \theta_s \left(m_2^{s1/2} z + m_1\right)\right) g \left(m_2^{s1/2} z + m_1\right) m_2^{s1/2}. \quad (30)$$

**Example 2:** Let’s consider the particular case of the truncated Normal GME density in Example 1, $\tilde{f}_N(x, \cdot)$. If we constrain the first two moments of $\tilde{f}_N(x, \cdot)$ to be equal to the first two moments of a truncated at the fourth moment EGC density, denoted as $\tilde{\pi}(x, d)$, (without loss of generality, we assume $m_1 = 0$ and $m_2 = 1$, i.e. $\gamma_1 = -3\gamma_3$, $\gamma_2 = -6\gamma_4$ and, $d_1 = d_2 = 0$), then the third and fourth moments of both densities must also be the same. As a consequence, $\tilde{f}_N(x, \gamma)$ and $\tilde{\pi}(x, d)$ are exactly the same distribution, which, furthermore, can be expressed in terms of the skewness ($sk = m_3^s/m_2^{s3/2}$) and kurtosis ($ku = m_4^s/m_2^{s2}$) coefficients as,

$$\tilde{f}_N(x, M) = \left[1 + \frac{sk}{3!} (x^3 - 3x) + \frac{ku - 3}{4!} (x^4 - 6x^2 + 3)\right] \phi(x). \quad (31)$$

**Proposition 6** A Positive GME pdf using a chosen pdf, $g(\cdot)$, as basis, denoted as $F(x, \gamma)$, can be obtained by squaring the polynomials of $f(x, \gamma)$ as,

$$F(x, \gamma) = \frac{1}{W} \left(1 + \sum_{s=1}^{n} \gamma_s^2 \Psi_s(x)\right) g(x), \quad (32)$$

where $W$ is the constant that guarantees that $F(x, \gamma)$ integrates to one,

$$W = \int \left(1 + \sum_{s=1}^{n} \gamma_s^2 \Psi_s(x)\right) g(x) dx = 1 + \sum_{s=1}^{n} \gamma_s^2 (\mu_{2s} - \mu_{2s}^2). \quad (33)$$

**Proposition 7** The non-central moments of $F(x, \gamma)$, denoted as $\tilde{m}_i$, can be expressed in terms of the moments of the density used as basis, as displayed in equation (34).

$$\tilde{m}_i = E[x^i] = \mu_i + \sum_{s=1}^{n} \gamma_s^2 \left[\mu_{2s+i} + \mu_s (\mu_i - 2\mu_{s+i})\right] \quad \forall i = 1, 2, 3, \ldots \quad (34)$$
Remark 1 Corollaries 1 and 2 straightforwardly apply to $F(x, \gamma)$, as a consequence of Proposition 7 and, because of the linear relation between the density moments and the squared parameters of that pdf.

Example 3: The standardized, symmetric ($\gamma_1 = \gamma_2 = 0$) and Positive Normal GME density expanded up to the fourth moment is given by,

$$F_N^*(x, \gamma) = \frac{1}{W^*} \left[ 1 + \gamma_2^2(cx^2 - 1)^2 + \gamma_4^2(c^2x^4 - 3)^2 \right] \phi(c^{1/2}x)c^{1/2},$$

where,

$$W^* = 1 + 2\gamma_2^2 + 96\gamma_4^2,$$

$$c = \frac{1 + 10\gamma_2^2 + 864\gamma_4^2}{W^*}.$$ (35)

Proposition 8 The cdf of a random variable $x \sim f_N(x, \gamma)$ can be computed as,

$$\Pr [x \leq a] = \int_{-\infty}^{a} f_N(x, \gamma)dx = \int_{-\infty}^{a} \phi(x)dx$$

$$+ \sum_{s=1}^{n} \frac{\gamma_s}{a^{s-1}} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \frac{\xi b^{s-3}}{a^{s-3}} \phi(a),$$ (37)

where,

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

Proposition 9 The cdf of a random variable $x \sim F_N(x, \gamma)$ is given by,

$$\Pr [x \leq a] = \int_{-\infty}^{a} F_N(x, \gamma)dx = \int_{-\infty}^{a} \phi(x)dx$$

$$+ \frac{2}{W} \sum_{s=1}^{n} \frac{\gamma_s^2}{a^{s-1}} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \frac{\zeta b}{a^{s-3}} \phi(a),$$

$$- \frac{1}{W} \sum_{s=1}^{n} \frac{\gamma_s^2}{a^{2s-1}} + (2s-1)a^{s-3} + (2s-1)(2s-3)a^{2s-5} + \ldots + \mu_2^s a^{3s} \phi(a),$$ (38)

where $W$ is the constant in equation (33) — for the moments of $\phi(x)$, denoted as $\mu_+^s$, $b$ is the constant defined above and,

$$\zeta = \begin{cases} (s-1)(s-3) \cdots 2 & \forall s \text{ odd} \\ (s-1)(s-3) \cdots 3a & \text{otherwise}. \end{cases}$$
Example 4: If $x$ is distributed according to $F^*_N(x, \gamma)$ — equation (35), the probability of any quantile $a$ is,

$$
\Pr [x \leq a] = \int_{-\infty}^{c^{1/2}a} F^*_N(x, \gamma)dx = \int_{-\infty}^{c^{1/2}a} \phi(x)dx + \frac{1}{W} 2\gamma_1^2 c^{1/2}a \phi(c^{1/2}a) \\
- \frac{1}{W}(\gamma_1^2 - 6\gamma_2^2) (c^{3/2}a^3 + 3c^{1/2}a) \phi(c^{1/2}a) + \\
- \frac{1}{W} \gamma_2^2 (c^{7/2}a^7 + 7c^{5/2}a^5 + 35c^{3/2}a^3 + 105c^{1/2}a^{1/2}) \phi(c^{1/2}a).
$$

(39)

3 Empirical application

In this section we study the applicability of the GME by means of a forecasting exercise for the conditional variance and VaR of the daily return on the British pound versus the US dollar (BP/$) exchange rate, $r_t$, over the period January 1983 to March 2002, for a total of $T = 4,882$ observations. The empirical distribution of $r_t$ presents a small negative skewness coefficient that is not statistically different from zero, which justifies the assumption of unconditionally symmetric distributions.

Let the conditional distribution of $r_t$, be either Gaussian, standardized Student’s t with $\nu$ degrees of freedom (Bollerslev, 1987), or Positive Normal GME (hereafter GME model), with conditional mean and variance following an AR(1) and a GARCH(1,1) process, respectively, i.e.,

$$
r_t = \xi_0 + \xi_1 r_{t-1} + u_t, 
$$

(40)

$$
u_t = h_t^{1/2} x_t, \ u_t|\Omega_{t-1} \sim N(0, h_t), \ u_t|\Omega_{t-1} \sim t_\nu(0, h_t), \ u_t|\Omega_{t-1} \sim GME(0, h_t),
$$

$$
h_t = \varphi_0 + \varphi_1 u_{t-1}^2 + \varphi_2 h_{t-1},
$$

(41)

where $\Omega_{t-1}$ denotes the econometrician information set up to time $t - 1$, and $h_t$ is the variance of the conditional distribution of $u_t$. The AR(1) process for the conditional mean was selected according to the Akaike Information Criterion (AIC). Note that the conditional Positive Normal GME pdf of $u_t$ is $h_t^{-1/2} F^*_N(x_t, \gamma)$, where $F^*_N(x_t, \gamma)$ is given in equation (35).
We use the first \( m = 4,381 \) observations to estimate the parameters of the models above, and compute \( N = 500 \) out-of-sample 1 step ahead forecasts of the conditional mean, \( \hat{\eta}_{t+1} \), and the conditional variance, \( \hat{\sigma}_{t+1} \), by using a rolling window of size \( m = T - N - 1 \) that discards old observations.

The estimation procedure is carried out, recursively, in two steps: Firstly, the AR(1) process is estimated by ordinary least squares. Secondly, the rest of parameters are estimated by (quasi)-maximum likelihood ((Q)ML) using the AR(1) residuals from the first step. Robust QML covariance estimators are calculated by means of Bollerslev and Wooldridge (1992) formula. The in-sample performance of the models is measured by using the mean of the AIC statistics over the \( N \) estimations. It is worth mentioning that the optimization of the GME model likelihood function was smoothly achieved providing that starting values were chosen adequately.

Table 1 contains the models estimation results. A first observation is that, there is a significant effect of the GME model on the sum of the GARCH coefficients; note that \( \hat{\varphi}_1 + \hat{\varphi}_2 \) is very near 1 in all three models but slightly lower in the case of the GME. In relation to the fit of the distribution tails, we obtain that the degrees of freedom coefficient, \( \hat{\nu} \), is around 5.6, even after correcting for volatility clustering, which together with the estimates of the polynomial weights in the GME model corroborates the existence of leptokurtosis in the returns conditional distribution. In relation to the in-sample goodness-of-fit, the Student’s t and the GME models provide an overall similar fit and both clearly outperform the Gaussian model.

[Table 1 Here]

Figure 1 shows an illustration of the models fitted in Table 1 in comparison to the histogram of our first data window. Observe that the GME model presents enough flexibility to capture not only the pick in the center of the distribution but also the density in the tails, while the Student’s t tends to overestimate the tails. This result is in line with those in Mauleón and Perote (1999), who also show a superior performance of semi-nonparametric density in relation to the Student’s t distribution.

[Figure 1 Here]
Figure 2 presents an illustration of the allowable shapes of the GME distribution depending on the values of its parameters. It is interesting to see how the density tails change allowing for heavy tails and multimodality.

3.1 Conditional variance forecasting performance

In this section we study the relative performance of the aforementioned models for forecasting the conditional variance of \( r_t \). For this purpose, the forecast loss is measured with respect to the out-of-sample squared residuals, \( \{ \hat{u}_t^2 \}_{t=m+1} \), by using the statistical loss functions family proposed by Patton (2006). So, the forecast at time \( t \), \( \hat{h}_{t+1} \), is compared with the realization \( \hat{u}_t^2 \), and the forecasting error is \( e_{t+1} = \hat{h}_{t+1} - \hat{u}_t^2 \). That class of loss functions is shown to be robust to models ranking when using imperfect volatility proxies, as e.g. the squared residuals, and includes: the squared error loss function, \( L_1 \), and asymmetric loss measures penalizing more heavily either under-predictions, \( L_2 \), or over-predictions, \( L_3 \).

\[
L_{1,t}(\hat{u}_t^2, \hat{h}_t) = e_t^2, \tag{42}
\]
\[
L_{2,t}(\hat{u}_t^2, \hat{h}_t) = \frac{\hat{u}_t^2}{\hat{h}_t} - \log(\hat{u}_t^2/\hat{h}_t) - 1, \tag{43}
\]
\[
L_{3,t}(\hat{u}_t^2, \hat{h}_t) = (\hat{u}_t^6 - \hat{h}_t^3)/6 - \hat{h}_t^2(\hat{u}_t^2 - \hat{h}_t)/2. \tag{44}
\]

The significance of the difference between these loss functions is tested by using the Diebold and Mariano (DM) (1995) test. For 1 step ahead forecast and a given loss function, \( L_j \), \( j = 1, 2, 3 \), the DM test null hypothesis of equal predictive ability of forecasts from two models \( f \) and \( g \) is,

\[
H_0 : E [d_{t+1}] = 0, \tag{45}
\]
\[
d_{t+1} = L_{j,t+1}(\hat{u}_{t+1}, \hat{h}_{t+1}^f) - L_{j,t+1}(\hat{u}_{t+1}, \hat{h}_{t+1}^g). \tag{46}
\]

The test statistic is, \( DM = \overline{d}/(2\pi \hat{\varphi}_d(\omega = 0)/N)^{1/2} \sim_a N(0, 1) \), where \( \overline{d} \) is the sample mean of the loss differential series over the out-of-sample period, and \( \hat{\varphi}_d(\omega = 0) \) is a consistent estimate of the loss differential spectral density function at frequency 0.

Table 2 presents the results of the DM test for all pairwise comparisons and loss functions. The entries are the means of the loss functions, \( L_j \), \( j = 1, 2, 3 \), over the out-of-sample period
for the models under comparison. The number within parentheses below each entry is the p-value of the test.

[Table 2 Here]

A sharp result that emerges from Table 2 is that there are not statistical differences between Gaussian and GME models but both significantly outperformed the Student’s t according to the squared error and error of underprediction loss functions, \( L_1 \) and \( L_3 \), respectively. The resulting model ranking is then: Gaussian \( \succeq \) GME \( \succ \) Student’s t. A second observation is that the Student’s t model tend to overpredict more and underpredicts less the volatility than the GME and the Gaussian models in this order, although differences are not statistically significant in relation to the error of overprediction loss function, \( L_2 \).

We note that, to some extent, this result may seem counterintuitive since one would expect a better performance from more parsimonious models such as, GME and Student’s t, in relation to the Gaussian, three possible explanations follow. First, the proxy we use for the unobservable "true" volatility, i.e. the daily squared residual, is too noisy and the use of a better proxy, for instance the realized volatility computed using intraday data, may give more reliable results, as suggested by Andersen and Bollerslev (1998) by using the \( R^2 \) of the Mincer-Zarnowitz regression. In contrast, Patton (2006) did not find differences in models rankings for forecasting volatility, by using loss functions with respect to those different proxies. These results suggest that models performance for volatility forecasting is very sensible to the statistical loss functions used, as argued in Hansen and Lunde (2005), among others. Further research in this line seems worthwhile. Second, the assumption on the distribution itself does not necessarily help to forecast asset returns volatility, unless the model is also capable of capturing the conditional variance dynamics (i.e. clustering, asymmetries, long-memory etc.) by means of an appropriate specification. Third, we argue that differences in performance between Gaussian, GME and Student’s t models (or other models assuming heavy-tailed distributions) for forecasting asset returns risk are more likely to appear when the aim is predicting VaR rather than volatility, in presence of leptokurtosis; see, for instance, Ñíguez (2008) for further research on these last two topics.
3.2 Tails forecasting performance

Our aim in this section is to comment on the models performance for forecasting the BP/$ exchange-rate return distribution tails. Thus, we compute $N$ 1 step ahead VaR forecasts for confidence levels $\alpha = \{0.1, 0.05, 0.025, 0.01\}$, denoted as $\{\widetilde{V aR}_t^\alpha\}_{t=m+1}^T$.

Following Engle (2001), the VaR forecasts corresponding to the estimated Gaussian conditional variance are calculated by using the percentile of the empirical distribution of the standardized residuals for every in-sample window, so that the model accounts for the observed excess kurtosis not captured when using standard Gaussian percentiles. For the sake of simplicity, the VaR forecasts obtained by means of this procedure are labelled under "Gaussian".

The models performance is assessed by using the following three criteria: unconditional coverage, $\hat{\alpha}$, the magnitude of the exception statistic, $M_\alpha$, (López, 1999), and the predictive quantile loss (PQL) function, $Q_\alpha$; see, for instance, Giacomini and Komunjer (2005) for further applications of PQL functions in a VaR forecasting context,

$$\hat{\alpha} = \frac{1}{N} \sum_{t=m+1}^{T} 1 \left\{ r_t < \widetilde{V aR}_t^\alpha \right\},$$

$$M_\alpha = \sum_{t=m+1}^{T} 1 \left\{ r_t < \widetilde{V aR}_t^\alpha \right\} \left( |r_t| - |\widetilde{V aR}_t^\alpha| \right)^2,$$

$$Q_\alpha = \frac{1}{N} \sum_{t=m+1}^{T} \left( \alpha - 1 \left\{ r_t < \widetilde{V aR}_{m,t}^\alpha \right\} \right) \left( r_t - \widetilde{V aR}_t^\alpha \right).$$

where $1 \{U\}$ is an indicator variable taking the value 1 if $U$ is true and 0 otherwise.

Table 3 presents the results of the VaR evaluation criteria. A first observation that emerges from this table is that, according to the unconditional coverage criterion all three models yield acceptable results for all significance levels considered. On the other hand, regarding the magnitude of the exception statistic and the PQL function, the GME model provides slightly more accurate VaR forecasts than the Student’s t for all significance levels, and than the Gaussian for 1%, 2.5% and 10% levels, the Gaussian being generally preferred to the Student’s t. A likely explanation for the good performance of the GME model may be the flexibility of the density expansions to parsimoniously fit the shape of the distribution.
4 Concluding remarks

Edgeworth and Gram-Charlier polynomial expansions have been usually applied to expand the Gaussian distribution. The advantages of the resulting EGC density are mainly related to its flexibility to adapt to small changes in the frequencies far in the distribution tails. On the other hand, those expansions present practical rigidities when the aim is to expand other distributions rather than the Normal, as argued in Mauleón and Perote (2000).

In this paper we have developed a new class of polynomial series expansions, that we name GME, which are based on the moments of an underlying distribution, chosen depending on the study case. We have analyzed their theoretical properties and have shown that GME preserve the flexibility characteristic of the Edgeworth and Gram-Charlier expansions but they are simpler and thus more tractable in some contexts. More specifically, we argue that the advantages of the densities based on GME in relation to EGC densities are the following: First, they possess a simpler polynomial structure that does not require orthogonality to prove statistical density properties as: positivity, up to one integration and quantile computation. Second, they are as much flexible or parsimonious as the EGC densities but with a more general formulation, namely, GME expansions can be more straightforwardly applied to expand any density, only requiring as many finite moments as the expansion order. In addition, we have particularly analyzed the GME of the Gaussian density and have provided the parameter restrictions under which Normal GME and EGC densities are equivalent.

We tested the performance of the GME densities by means of an empirical application for forecasting the conditional variance and VaR of BP/$ exchange-rate returns, comparing the GME model with Gaussian and Student’s t models. The forecasts were evaluated by using
the class of statistical loss functions proposed by Patton (2006), the unconditional coverage
and magnitude of the exceptions statistics in López (1999), and the PQL function of Koenker
and Bassett (1978). In summary, our empirical results show that the GME model is as good
as the Gaussian and both are significantly better than the Student’s t for forecasting the
volatility. On the other hand, when the aim is forecasting VaR, we find evidence of that the
GME model provides more accurate forecasts in relation to the Student’s t and the Engle’s
Appendix: Proofs of propositions

Proof of Proposition 1:
The GME density integrates to one:
\[
\int f(x, \gamma) dx = \int g(x) dx + \sum_{s=1}^{q} \gamma_s \int (x^s - \mu_s) g(x) dx = 1 + 0 = 1.
\]

Proof of Proposition 2:
The GME density defined in equation (13) can be rewritten as,
\[
f(x, \gamma) = \left( \sum_{s=1}^{n} \gamma_s x^s + k \right) g(x)
\]
where \( k = 1 - \sum_{s=1}^{n} \gamma_s \mu_s \). Therefore, if \( 0 \leq \gamma_s \leq \frac{1}{\mu_s} \) \( \forall s \) even, and \( \gamma_s = 0 \) \( \forall s \) odd, then it is clear that \( \sum_{s=1}^{n} \gamma_s x^s + k \geq 0 \), since \( \sum_{s=1}^{n} \gamma_s \mu_s \leq 1 \). Consequently, \( f(x, \gamma) \geq 0 \).

Proof of Proposition 3:
The GME density non-central moments can be straightforwardly obtained through the standard Normal density moments:
\[
E[x^i] = \int x^i f(x, \gamma) dx
\]
\[
= \int x^i g(x) dx + \sum_{s=1}^{n} \gamma_s \int x^i (x^s - \mu_s) g(x) dx = \mu_i + \sum_{s=1}^{n} \gamma_s (\mu_{s+i} - \mu_s \mu_i).
\]

Proof of Proposition 4:
The first \( n - th \) moments of the EGC density can be obtained from equations (9) and (11) as follows:
\[
\text{M}_{EGC} = E[Z] = \int Z (1 + H^T d) \phi(x) dx
\]
\[
= \int B^{-1}(H - \Gamma^T \mu^+) (1 + H^T d) \phi(x) dx
\]
\[
= B^{-1} \int H \phi(x) dx
\]
\[
+ B^{-1} \int HH^T d \phi(x) dx - B^{-1} \Gamma^T \mu^+ \int \phi(x) dx - B^{-1} \Gamma^T \mu^+ \int H^T d \phi(x) dx
\]
\[
= 0 + B^{-1} (S \times d) - B^{-1} \Gamma^T \mu^+ + 0
\]
\[
= B^{-1} [(S \times d) - \Gamma^T \mu^+].
\]
On the other hand, the first $n-th$ moments of the Normal GME, $f_N(x; \gamma)$, can be expressed as,

$$M_{GME} = A\gamma + \mu^+.$$ 

Therefore, $M_{GME} = M_{EGC}$ if, and only if, $A\gamma + \mu^+ = B^{-1} [(S \times d) - I'\mu^+]$ if, and only if, $\gamma = A^{-1}B^{-1}Sd - A^{-1}(B^{-1} + I')\mu$. 

**Proof of Proposition 5:**

If $x \sim f(x; \cdot) —$ equation (13) or (21), then the standardized variable (first and second moments zero and one, respectively) $z = \frac{x - m_1}{m_2^{1/2}} \sim f^*(z; \cdot) —$ equations (29) or (30), respectively.

**Proof of Proposition 6:**

The constant $W$ that makes $F(x, \gamma) —$ equation (32), to integrate to one can be obtained as,

$$W = \int \left(1 + \sum_{s=1}^{n} \gamma_s^2(x^s - \mu_s)^2\right) g(x)dx$$

$$= \int g(x)dx + \sum_{s=1}^{n} \gamma_s^2 \left( \int x^{2s}g(x)dx + \mu_s^2 \int g(x)dx - 2\mu_s \int x^s g(x)dx \right)$$

$$= 1 + \sum_{s=1}^{n} \gamma_s^2 (\mu_{2s} + \mu_s^2 - 2\mu_s^2) = 1 + \sum_{s=1}^{n} \gamma_s^2 (\mu_{2s} - \mu_s^2).$$

**Proof of Proposition 7:**

The non-central moments of $F(x, \gamma) —$ equation (32), can also be formulated in terms of the moments of the density chosen as basis, $g(x)$:

$$E[x^i] = \int x^i \left(1 + \sum_{s=1}^{n} \gamma_s^2(x^s - \mu_s)^2\right) g(x)dx$$

$$= \int x^i g(x)dx$$

$$+ \sum_{s=1}^{n} \gamma_s^2 \left( \int x^{2s+i}g(x)dx + \mu_s^2 \int x^i g(x)dx - 2\mu_s \int x^{s+i} g(x)dx \right)$$

$$= \mu_i + \sum_{s=1}^{n} \gamma_s^2 (\mu_{2s+i} + \mu_s^2 \mu_i - 2\mu_s\mu_{s+i})$$

$$= \mu_i + \sum_{s=1}^{n} \gamma_s^2 \left[\mu_{2s+i} + \mu_s (\mu_i - 2\mu_{s+i})\right].$$
Proof of Proposition 8:

The cdf of a random variable $x$ distributed according to $f_N(x, \gamma) —$ equation (22), is:

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

$$= \int_{-\infty}^{a} \phi(x)dx$$

$$- \sum_{s=1}^{n} \gamma_s \mu_s^+ \int_{-\infty}^{a} \phi(x)dx + \sum_{s=1}^{n} \gamma_s \int_{-\infty}^{a} x^s \phi(x)dx + \sum_{s=1}^{n} \gamma_s \int_{-\infty}^{a} x^s \phi(x)dx$$

$$= \int_{-\infty}^{a} \phi(x)dx$$

$$- \sum_{s=1}^{n} \gamma_s \left[(x^{s-1} + (s - 1)x^{s-3} + (s - 1)(s - 5)x^{s-5} + \ldots + \xi x)b\right]\phi(x)$$

$$= \int_{-\infty}^{a} \phi(x)dx - \sum_{s=1}^{n} \gamma_s \left(a^{s-1} + (s - 1)a^{s-3} + (s - 1)(s - 3)a^{s-5} + \ldots + \xi a\right)\phi(a),$$

where,

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

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

$$\int x^s g(x)dx = \int x^{s-1} xg(x)dx = -x^{s-1} g(x) + (s - 1) \int x^{s-2} \phi(x)dx$$

since,

$$u = x^{s-1} \Rightarrow du = (s - 1)x^{s-2}dx,$$

$$dv = xg(x)dx \Rightarrow v = \int x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}dx = -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} = -\phi(x).$$

Therefore, by repeating the same argument recursively,

$$\int x^s \phi(x)dx = \begin{cases} 
- [x^{s-1} + (s - 1)x^{s-3} + (s - 1)(s - 3)x^{s-5} + \ldots + \xi]\phi(x) & \forall s \text{ odd} \\
\mu_s^+ \int \phi(x)dx - (x^{s-1} + (s - 1)x^{s-3} + (s - 1)(s - 3)x^{s-5} + \ldots + \xi x)\phi(x) & \text{otherwise}
\end{cases}$$

where $\mu_s^+ = \xi$. Furthermore, by applying recursively the l’Hôpital rule it is obtained,

$$\lim_{x \to -\infty} [x^s \phi(x)] = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} x^s = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} x^{-\frac{s}{2}} = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}}$$

$$= \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{s(s-2)\ldots}{x^{s-2}}} = \lim_{x \to -\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{s(s-2)\ldots}{x^{s-2}}} = 0.$$

This completes the proof.  

Proof of Proposition 9:

If $x$ is a random variable distributed according to the Positive Normal GME, $F_N(x, \gamma)$ — equation (32), its cdf can be obtained as follows:

$$
\int_{-\infty}^{a} F_N(x, \gamma) dx = \frac{1}{W} \int_{-\infty}^{a} \left( 1 + \sum_{s=1}^{n} \gamma_s^2 (x^s - \mu_s^+)^2 \right) \phi(x) dx
$$

$$
= \frac{1}{W} \int_{-\infty}^{a} \left( 1 + \sum_{s=1}^{n} \gamma_s^2 \mu_s^{+2} + \sum_{s=1}^{n} \gamma_s^2 x^{2s} - 2 \sum_{s=1}^{n} \gamma_s^2 \mu_s^+ x^s \right) \phi(x) dx
$$

$$
= \frac{1}{W} \int_{-\infty}^{a} \phi(x) dx + \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_s^{+2} \int_{-\infty}^{a} \phi(x) dx + \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \int_{-\infty}^{a} x^{2s} \phi(x) dx
$$

$$
- 2 \frac{1}{W} \sum_{s=1}^{n} \gamma_s \mu_s^+ \int_{-\infty}^{a} x^s \phi(x) dx - 2 \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_s^+ \int_{-\infty}^{a} x^s \phi(x) dx
$$

$$
= \frac{1}{W} \int_{-\infty}^{a} \phi(x) dx + \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_s^{+2} \int_{-\infty}^{a} \phi(x) dx
$$

$$
- \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \left[ (x^{2s-1} + (2s-1)x^{2s-3} + (2s-1)(2s-3)x^{2s-5} + \ldots + \mu_2^+ x \phi(x)_{-\infty}^a \right]
$$

$$
+ \frac{2}{W} \sum_{s=1}^{n} \gamma_s^2 \left[ (x^{s-1} + (s-1)x^{s-3} + (s-1)(s-3)x^{s-5} + \ldots + \zeta^{s} x \phi(x)_{-\infty}^a \right]
$$

$$
+ \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_2^{+} \int_{-\infty}^{a} \phi(x) dx - 2 \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_s^+ \int_{-\infty}^{a} \phi(x) dx
$$

$$
= \int_{-\infty}^{a} \phi(x) dx
$$

$$
+ \frac{2}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_s^+ \left[ (x^{s-1} + (s-1)x^{s-3} + (s-1)(s-3)x^{s-5} + \ldots + \zeta^{s} x \phi(x)_{-\infty}^a \right]
$$

$$
- \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \left[ (x^{2s-1} + (2s-1)x^{2s-3} + (2s-1)(2s-3)x^{2s-5} + \ldots + \mu_2^+ x \phi(x)_{-\infty}^a \right]
$$

$$
= \int_{-\infty}^{a} \phi(x) dx
$$

$$
+ 2 \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 \mu_s^+ (a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \zeta^{s} a \phi(a)
$$

$$
- \frac{1}{W} \sum_{s=1}^{n} \gamma_s^2 (a^{2s-1} + (2s-1)a^{2s-3} + (s-1)(s-3)a^{2s-5} + \ldots + \mu_2^+ a) \phi(a)
$$

where,

$$
\zeta = \begin{cases} 
  (s-1)(s-3) \cdots 2 & \text{if } s \text{ odd} \\
  (s-1)(s-3) \cdots 3a & \text{otherwise}
\end{cases} \quad \text{and } b = \begin{cases} 
  1 & \text{if } s \text{ even} \\
  0 & \text{otherwise}
\end{cases}
$$
### TABLE 1
Estimation results

#### Mean equation:
\[ r_t = \phi_0 + \phi_1 r_{t-1} + u_t, \quad u_t = h_t^{1/2} x_t \]
\[ u_t|\Omega_{t-1} \sim N(0, h_t), \quad u_t|\Omega_{t-1} \sim GME(0, h_t), \quad u_t|\Omega_{t-1} \sim t_\nu(0, h_t) \]

#### Variance equation:
\[ h_t = \varphi_0 + \varphi_1 u_{t-1}^2 + \varphi_2 h_{t-1} \]

<table>
<thead>
<tr>
<th></th>
<th>Parameters</th>
<th>Gaussian</th>
<th>GME</th>
<th>Student’s t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean equation</td>
<td>$\xi_0$</td>
<td>0.0014</td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi_1$</td>
<td></td>
<td>0.0778</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td></td>
<td>(5.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance equation</td>
<td>$\varphi_0$</td>
<td>0.0030</td>
<td>0.0020</td>
<td>0.0024</td>
</tr>
<tr>
<td>(t Stat)</td>
<td>(3.66)</td>
<td>(3.19)</td>
<td>(2.43)</td>
<td></td>
</tr>
<tr>
<td>$\varphi_1$</td>
<td>0.0432</td>
<td>0.0339</td>
<td>0.0493</td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td>(8.22)</td>
<td>(7.66)</td>
<td>(6.26)</td>
<td></td>
</tr>
<tr>
<td>$\varphi_2$</td>
<td>0.9494</td>
<td>0.9459</td>
<td>0.9467</td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td>(150.2)</td>
<td>(130.7)</td>
<td>(110.5)</td>
<td></td>
</tr>
<tr>
<td>Weights</td>
<td>$\gamma_2$</td>
<td>0.0966</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td>(2.82)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_4$</td>
<td>-0.0215</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td>(-11.45)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DoF</td>
<td>$\nu$</td>
<td>5.662</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(t Stat)</td>
<td>(11.29)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>1.7910</td>
<td>1.7514</td>
<td>1.7433</td>
<td></td>
</tr>
</tbody>
</table>

The reported coefficients presented in this table are (Q)ML estimates of the AR(1)-GARCH(1,1) processes under the Gaussian, the Student’s t or the GME distributions, for the BP/$ exchange-rate daily returns. $\gamma_s$ denotes the weight parameter of the $s$ -th order polynomial in the GME distribution. DoF denotes degrees of freedom, and AIC is the mean of the AICs of the $N$ estimations through the out-of-sample period. t Statistics calculated from robust standard errors are in parentheses below the parameter estimates.
<table>
<thead>
<tr>
<th>Models</th>
<th>Gaussian</th>
<th>GME</th>
<th>Student’s t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$L_1$</td>
<td></td>
</tr>
<tr>
<td>GME</td>
<td>0.15822</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.200)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student’s t</td>
<td>0.15847</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td></td>
<td>0.15905</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.048)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$L_2$</td>
<td></td>
</tr>
<tr>
<td>GME</td>
<td>-0.41228</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.207)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student’s t</td>
<td>-0.41266</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.312)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td></td>
<td>-0.41141</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.714)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$L_3$</td>
<td></td>
</tr>
<tr>
<td>GME</td>
<td>0.05586</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student’s t</td>
<td>0.05593</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td></td>
<td>0.05606</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.012)</td>
<td></td>
</tr>
</tbody>
</table>

This table contains the results of the DM predictive ability test for the models and loss functions presented in Section 3. The entries are the means of the loss functions $L_j$, $j = 1, 2, 3$ over the out-of-sample period for the models in the columns. The numbers within parentheses are DM test t-statistics for the predictive ability of the model in the column versus the model in the row under the loss function $L_j$, $j = 1, 2, 3$, over the out-of-sample period. Predictions 500.
This table contains the results of the VaR tests described in Section 3.2. $\hat{\alpha}$ denotes the estimated unconditional coverage probability, $C_\alpha$ denotes the magnitude of the exception statistic and, $Q_\alpha$ is the predictive quantile loss function, for 1 step ahead VaR forecasts with significance levels $\alpha = 0.1, 0.05, 0.025, 0.01$, obtained with the Gaussian, GME and Student’s t models. Predictions 500.
Panel A shows the histogram and fitted distribution of the in-sample returns from Gaussian, Student’s t and GME models. Panel B highlights the fit of the left tail.
The Figure shows the allowable shape of the density in terms of the values of its parameters, p2 and p3 in the figure correspond to $\gamma_2$ and $\gamma_4$, respectively, in text.
Plots of 10%, 5%, 2.5% and 1% VaR forecasts from the GME model against out-of-sample BP/$ exchange-rate returns. Predictions 500.
References


