WESTMINSTER
BUSINESS SCHOOL

WORKING PAPER SERIES
IN BUSINESS AND MANAGEMENT

WORKING PAPER 14-1

February 2014

The Moment Expansions: A Semi-nonparametric Method with Applications for Risk Management

Trino-Manuel Niguez* and Javier Perote[+]

* *Westminster Business School*
+ *University of Salamanca*

Corresponding author:

Trino-Manuel Niguez
Westminster Business School
University of Westminster
35 Marylebone Road
London NW1 5LS, UK
T.M.Niguez@westminster.ac.uk

ISBN ONLINE: 978-1-908440-19-8

http://www.westminster.ac.uk/about-us/schools/business/research/working-paper-series

UNIVERSITY OF
WESTMINSTER

The Moment Expansions: A Semi-nonparametric Method with Applications for Risk Management

Trino-Manuel Ñíguez

Department of Economics and Quantitative Methods, Westminster Business School, University of Westminster, London NW1 5LS, UK

Javier Perote

Department of Economics, University of Salamanca, 37007 Salamanca, Spain

Abstract

This paper presents a novel family of semi-nonparametric (SNP) distributions whose polynomials are defined as the difference of the n-th power of the variable and the n-th moment of the density being expanded. We show that the so-obtained Moment Expansions (ME) pdf exhibits empirical and theoretical advantages derived from its simple and general specification that make it a useful alternative to existing SNP pdfs. We test the applicability of our approach through a comparative empirical application for forecasting financial risk. We show that a Normal-ME model presents a relatively good forecasting peformance that together with its statistical features makes it a useful methodlogy in risk management.

Keywords: Gram-Charlier Series; Semi-nonparametric methods; Value-at-Risk.

JEL Classification numbers: C53, G12.

1 Introduction

During the last decades, the literature related to the modelling and forecasting of financial variables asymmetric and heavy-tailed distribution has undergone a huge development. Among the existing methodologies, the semi-nonparametric (SNP hereafter) techniques have been proven very useful for that purpose; see, for instance, Vilhelmsson (2009) and Del Brio et al. (2011) for recent SNP applications for forecasting financial risk. SNP probability density functions (pdf hereafter) feature a flexible specification that allows to approximate any "true" target distribution at any desired degree of accuracy; a result that stems from the seminal papers of Edgeworth (1896, 1907) and Charlier (1905). In particular, a frequency function can be expanded in an (infinite) series of derivatives of a Normal pdf which gives rise to the so-called Edgeworth and Gram-Charlier (GC hereafter) pdf. GC (Type A) series were brought into econometrics by Sargan (1975, 1976), and later extensively developed by authors such as: Jarrow and Rudd (1982), Gallant and Nychka (1987), Gallant and Tauchen (1989), Corrado and Su (1997), Mauleón and Perote (2000), Jondeau and Rockinger (2001), Velasco and Robinson (2001), Verhoeven and McAleer (2004), León et al. (2005, 2009), Ñíguez and Perote (2012) and Ñíguez et al. (2012), among others. These papers provide analyses of the SNP densities' theoretical properties and their wide variety of applications, which range from hypothesis testing to portfolio choice and financial risk forecasting.

The practical application of SNP pdfs require the truncation of its polynomial expansion, the resulting truncated function is not really a pdf since it may yield negative values for subsets of its parametric space. This well-known definitional issue, firstly highlighted by Barton and Dennis (1952), has been tackled in the literature in different ways: (*i*) through parametric restrictions (Jondeau and Rockinger 2001), (*ii*) through monitored optimization (Mauleón and Perote 2000), and (*iii*) through density reformulations based on the methodology of Gallant and Nychka (1987) and Gallant and Tauchen (1988), see e.g. León et al. (2009) and Ñíguez and Perote (2012). Another practical and theoretical issues typical of SNP pdfs are: probable maximum-likelihood sub-optimization associated with multimodality, and difficult implementation for expansions of non-Gaussian pdfs.

This paper addresses these issues by proposing a new type of polynomial expansion that we name Moment Expansions (ME hereafter). We analyze the ME theoretical properties and show that ME pdfs are very theoretically tractable and easy to use in practice. With particular regard to the well-known definitional issue of GC pdfs, we show that well-defined (positive in the whole parametric space) ME pdfs are more flexible and easier to implement as: (i) they involve weaker restrictions on the density parameters and (*ii*) they preserve linearity between density moments and parameters for Gallant and Nychka's (1987) type of positivity transformations. We test our proposed method through an empirical application for forecasting financial risk. The out-of-sample forecasting performance of a ME of the Normal pdf (Normal-ME hereafter) pdf with respect to Gaussian and Student's t pdfs is assessed through the following criteria: ranking-robust loss functions for imperfect volatility proxies (Patton, 2010); Value-at-Risk (VaR hereafter) predictive accuracy criteria (López, 1999); and the predictive quantile loss function (Koenker and Bassett, 1978). Our results show that: (*i*) Normal-ME and Gaussian pdfs provide a similar forecasting performance for the conditional variance, both being superior to the Student's t pdf, and (*ii*) the ME model yields more accurate VaR forecasts than the Gaussian-VaR method of Engle (2001), and the Student's t model.

2 The Moment Expansions

In this section, we define the ME and analyse its statistical properties. In order to set up notation we first summarize some results of the SNP literature. Let a random variable x be Gram-Charlier distributed with pdf given by,

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

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

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

It is well-known that the function in equation (1) is not really a well-defined density since it may yield negative values. Recent methods proposed in the literature, based in the methodology of Gallant and Nychka (1987) and Gallant and Tauchen (1988), to ensure the positiveness of pdfs based on Hermitian expansions include, the GC pdf in León et al. (2005, 2009) and the Positive Edgeworth-Sargan pdf in Ñíguez and Perote (2012). $\pi(x, \mathbf{d})$ can be re-written in matrix form as,

$$\pi(x, \mathbf{d}) = (1 + \mathbf{H}'\mathbf{d})\,\phi(x). \tag{3}$$

$$\mathbf{H} = \mathbf{BZ} + \mathbf{I}^*\boldsymbol{\mu}^+ \tag{4}$$

where $\mathbf{B}$, equation (5), is a lower triangular matrix of order n containing the coefficients of x^n in $\mathbf{H}$, $\mathbf{Z} = (x, x^2, \cdots, x^n)' \in \mathbb{R}^n$, $\boldsymbol{\mu}^+ = \left(\mu_1^+, \mu_2^+, \cdots, \mu_n^+\right)' \in \mathbb{R}^n$ is a vector containing the first $n-th$ order moments of $\phi(\cdot)$, $r = n/2$, and $\mathbf{I}^* = diag\,\{0, -1, 0, 1, 0, -1, \cdots, (-1)^r\}$ is a diagonal matrix of order n that includes the sign of the corresponding intercept of every Hermite polynomial.[1]

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ -3 & 0 & 1 & 0 & \cdots & 0 \\ 0 & -6 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dfrac{(-1)^{r-1}n!}{2^{r-1}(r-1)!2!} & 0 & \dfrac{(-1)^{r-2}n!}{2^{r-2}(r-2)!4!} & \cdots & 1 \end{bmatrix} \tag{5}$$

GC expansions are defined for any continuous and differentiable parametric pdf. However, expansions of non-Normal pdfs result in rather complex specifications that difficult its

[1] Without loss of generality, we define the matrices $\mathbf{B}$ and $\mathbf{I}^*$, and the vector $\boldsymbol{\mu}^+$ for n even.

use for empirical and theoretical analyses. This complexity accentuates particularly when guaranteeing positivity is needed since Gallant and Nychka (1987) type of re-formulations yield non-linear relations between moments and parameters. In this paper, we propose the ME as a feasible solution to these issues. ME are defined in terms of the moments of the distribution being expanded and are valid to apply to any parametric pdf, only requiring the expanded pdf has finite moments up to the truncation order n.

Definition 1 *A ME of a pdf, $g(\cdot)$, with finite non-central moments up to the truncation order n, $E[x^s] = \mu_s \; \forall s = 1, 2, \ldots, n$, is defined as,*

$$f(x, \boldsymbol{\gamma}) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s(x)\right) g(x), \tag{6}$$

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

$$\Psi_s(x) = x^s - \mu_s. \tag{7}$$

Equation (6) can be re-written in matrix form as,

$$f(x, \boldsymbol{\gamma}) = (1 + (\mathbf{Z} - \boldsymbol{\mu})'\boldsymbol{\gamma})\, g(x). \tag{8}$$

where $\boldsymbol{\mu} \in \mathbb{R}^n$ contains the n-th order moments of $g(\cdot)$.

2.1 ME statistical properties

In this section we discuss the ME main properties (enumerated below), which are formally presented in Propositions 1-9 and Corollaries 1 and 2 in the Appendix, for the sake of clarity in the exposition.

1. ME pdfs integrate up to one (Proposition 1).

2. ME pdfs are positive in a wide range of the parameter space (Proposition 2).

3. The moment generating function of a ME pdf is straightforwardly obtained from the moments of the expanded pdf (Proposition 3).

4. The GC pdf can be obtained as a particular case of the Normal-ME pdf (Proposition 4).

5. Standardized ME pdfs are obtained through a linear transformation (Proposition 5).

6. ME pdfs admits Gallant and Nychka's (1987) type of transformations to ensure positiveness (Proposition 6).

7. The moments of 'positive' Normal-ME pdfs are linear functions of the squared density parameters (Proposition 7).

8. Closed forms for the cdf of Normal-ME and its transformed positive version can be straightforwardly derived (Propositions 8 and 9).

9. If a symmetric density is used as basis, then the ME even/odd moments depend exclusively on its even/odd parameters (Corollary 1).

10. The ME pdf can be expressed alternatively in terms of its moments (Corollary 2).

These properties pose the ME as a novel methodology for deriving SNP densities. In particular, the case of a Normal-ME pdf deserves especial attention as it inherits the good asymptotic properties of the GC approximation, thus being a good alternative to the latter for modelling the salient features of financial and economic variables. Examples 1 and 2 below provide a discussion of the Normal-ME and its features.[2]

[2]It is worth noting that Corollaries 1 and 2 apply to 'positive' Normal-ME but do not to 'positive' GC. Besides, feature 6 does not hold for positive transformations since 'positive' GC pdfs involve nonlinearity between density moments and parameters (see also Remark 1).

Example 1 A Normal-ME pdf, $f_N(\cdot)$, is defined as

$$f_N(x, \boldsymbol{\gamma}) = \left(1 + \sum_{s=1}^{n} \gamma_s (x^s - \mu_s^+)\right) \phi(x). \tag{9}$$

where

$$\mu_s^+ = \begin{cases} \dfrac{s!}{2^{\frac{s}{2}}(s/2)!} = (s-1)(s-3)(s-5)\cdots 3, & \forall s \text{ even} \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

From Proposition 2 (in the Appendix), we can ensure that this density is well-defined (positiveness is guaranteed) provided that $0 \leq \gamma_s \leq \dfrac{2^{\frac{s}{2}}(s/2)!}{s!}$ $\forall s$ even, and $\gamma_s = 0$ $\forall s$ odd. Furthermore, its moments can easily be computed as a direct application of Proposition 3 as,

$$E[x^i] = \mu_i^+ + \sum_{s=1}^{n} \gamma_s (\mu_{s+i}^+ - \mu_s^+ \mu_i^+) = \frac{i!}{\frac{i}{2}!} \frac{1}{2^{\frac{i}{2}}} \left[1 + \sum_{s=1}^{n} \gamma_s \frac{\frac{i}{2}!}{i!} \frac{1}{2^{\frac{s}{2}}} \left(\frac{(s+i)!}{\frac{(s+i)}{2}!} - \frac{i!s!}{\frac{i}{2}!\frac{s}{2}!} \right) \right]. \tag{11}$$

Example 2 A Normal-ME pdf truncated at $n = 4$ is given by

$$\widetilde{f}_N(x, \boldsymbol{\gamma}) = \left[1 + \gamma_1 x + \gamma_2 (x^2 - 1) + \gamma_3 x^3 + \gamma_4 (x^4 - 3)\right] \phi(x), \tag{12}$$

which can alternatively be re-written in terms of its first four moments as,

$$\widetilde{f}_N(x, \mathbf{M}) = [1 + m_1 \theta_1(x) + (m_2 - 1)\theta_2(x) + m_3 \theta_3(x) + (m_4 - 3)\theta_4(x)] \phi(x), \tag{13}$$

where

$$\theta_i(x) = \begin{cases} x(5 - x^2)/2, & if\ i = 1, \\ 2(x^2 - 1) - (x^4 - 3)/4, & if\ i = 2, \\ x(x^2 + 3)/6, & if\ i = 3, \\ (x^4 - 3)/24 - (x^2 - 1)/4, & if\ i = 4. \end{cases} \tag{14}$$

As a result, if we constrain the first two moments of $\widetilde{f}_N(x, \cdot)$ to be equal to the first two moments of a GC pdf truncated at $n = 4$, denoted as $\widetilde{\pi}(x, d)$, then the third and fourth-order moments of both densities must also be the same.[3] As a consequence, $\widetilde{f}_N(x,\gamma)$ and $\widetilde{\pi}(x,d)$ are exactly the same distribution (see Proposition 4), which 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,[4]

$$\widetilde{f}_N(x,\mathbf{M}) = \left[1 + \frac{sk}{3!}(x^3 - 3x) + \frac{ku-3}{4!}(x^4 - 6x^2 + 3)\right]\phi(x). \tag{15}$$

Furthermore, a standardized symmetric ($\gamma_1 = \gamma_3 = 0$) and positive Normal-ME density truncated at $n = 4$ is defined as,

$$F_N^*(x,\boldsymbol{\gamma}) = \frac{1}{W^*}\left[1 + \gamma_2^2(cx^2 - 1)^2 + \gamma_4^2(c^2x^4 - 3)^2\right]\phi(c^{1/2}x)c^{1/2}, \tag{16}$$

where $W^* = 1 + 2\gamma_2^2 + 96\gamma_4^2$ and $c = \frac{1+10\gamma_2^2+864\gamma_4^2}{W^*}$ (see Proposition 5 in the Appendix). The probability of any quantile a of $F_N^*(x,\gamma)$ can be obtained as, (see Propositions 8 and 9),

$$\begin{aligned}
\Pr\left[x \le a\right] = \int_{-\infty}^{c^{1/2}a} F_N^*(x,\boldsymbol{\gamma})dx = \int_{-\infty}^{c^{1/2}a} \phi(x)dx + \frac{1}{W}2\gamma_1^2 c^{1/2}a\phi(c^{1/2}a) \\
- \frac{1}{W}(\gamma_1^2 - 6\gamma_2^2)\left(c^{3/2}a^3 + 3c^{1/2}a\right)\phi(c^{1/2}a) + \\
- \frac{1}{W}\gamma_2^2\left(c^{7/2}a^7 + 7c^{5/2}a^5 + 35c^{3/2}a^3 + 105c^{1/2}a^{1/2}\right)\phi(c^{1/2}a).
\end{aligned} \tag{17}$$

3 Empirical application

This section provides an analysis of the applicability of the ME 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.

Let the conditional distribution of r_t, be either Gaussian, standardized Student's t with ν degrees of freedom (Bollerslev, 1987), or Positive Normal-ME (hereafter N-ME$^+$), with

[3] 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 $d_1 = d_2 = 0$.

[4] It is worth noting that equation (15) is the traditionally employed GC density in financial applications; see e.g. Jarrow and Rudd (1982), Jondeau and Rockinger (2001) or León et al. (2005).

conditional mean and variance following AR(1) (selected according to the Akaike Information Criterion (AIC)) and GARCH(1,1) processes, respectively, i.e.,

$$
\begin{aligned}
r_t &= \delta_0 + \delta_1 r_{t-1} + u_t, && (18)\\
u_t &= h_t^{\frac{1}{2}} x_t,\ u_t|\Omega_{t-1} \sim N(0, h_t),\ u_t|\Omega_{t-1} \sim t_\nu(0, h_t),\ u_t|\Omega_{t-1} \sim N - ME^{+}(0, h_t), &&\\
h_t &= \varphi_0 + \varphi_1 u_{t-1}^2 + \varphi_2 h_{t-1}, && (19)
\end{aligned}
$$

where Ω_{t-1} denotes the information set up to time $t-1$, and h_t is the variance of the conditional distribution of u_t.[5] We use the first $T - R - 1 = 4,381$ observations as the first in-sample window, and compute $R = 500$ out-of-sample 1-step-ahead forecasts of the conditional mean, $\widehat{r}_{t+1} = \widehat{E}(r_{t+1})$ and the conditional variance, $\widehat{h}_{t+1}$, by using a rolling window that discards old observations. The models are estimated by quasi-maximum likelihood (QML), the covariance estimates are robust Bollerslev and Wooldridge (1992). Table 1 contains the estimation results.[6] We observe an effect of the N-ME$^+$ on the sum of the GARCH coefficients; $\widehat{\varphi}_1 + \widehat{\varphi}_2$ is near 1 in all three models but lower in the case of the N-ME$^+$. The degrees of freedom coefficient, $\widehat{\nu}$, is around 5.6, which together with the estimates of the N-ME$^+$ model shows that there is leptokurtosis in the returns distribution. The Student's t and the N-ME$^+$ provide a similar goodness-of-fit and both outperform the Gaussian, according to the mean of the AIC statistics over the R estimations. Figure 1 shows an illustration of the models fitted in Table 1. The N-ME$^+$ captures the sharp peak in the center of the distribution and the tails shape, while the Student's t seems to overestimate the tails. Figure 2 presents an illustration of the shapes of the N-ME$^+$ pdf, it is interesting to observe how the density shape responds to changes in the values of the parameters allowing for heavy tails and multimodality.

[5] Note that the conditional N-ME$^+$ pdf of u_t is $h_t^{-\frac{1}{2}} F_N^*(x_t, \gamma)$, where $F_N^*(x_t, \gamma)$ is given in equation (16).

[6] It is worth noting that the optimization of the N-ME$^+$ model likelihood function was smoothly achieved after starting values were chosen adequately.

TABLE 1

Estimation results

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

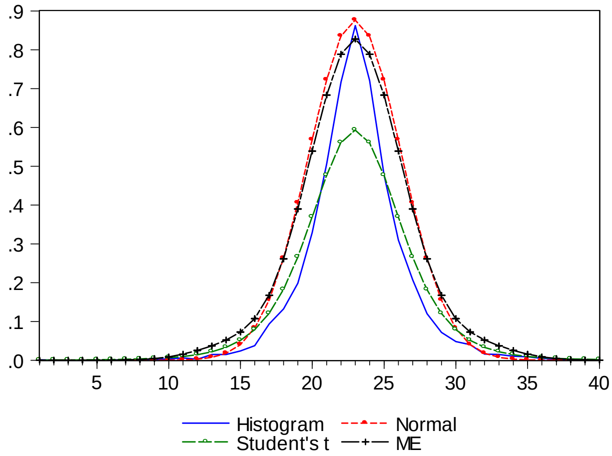
$u_t|\Omega_{t-1} \sim N(0, h_t), \quad u_t|\Omega_{t-1} \sim N\text{-}ME^{+}(0, h_t), \quad u_t|\Omega_{t-1} \sim t_{\nu}(0, h_t)$

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

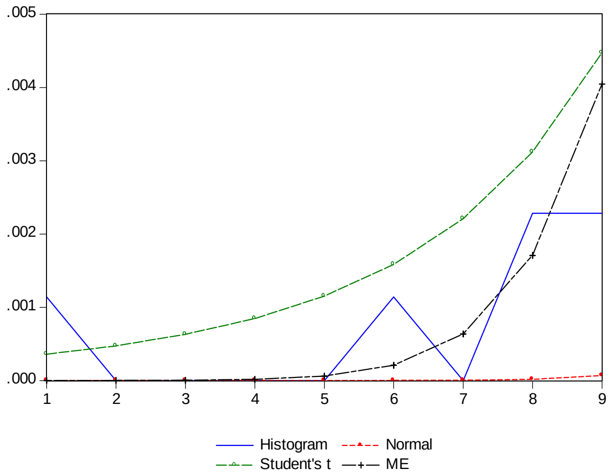
	Parameters	Gaussian	N-ME^{+}	Student's t
Mean equation	δ_0		0.0014	
			(0.12)	
	δ_1		0.0778	
			(5.16)	
Variance equation	φ_0	0.0030	0.0020	0.0024
		(3.66)	(3.19)	(2.43)
	φ_1	0.0432	0.0339	0.0493
		(8.22)	(7.66)	(6.26)
	φ_2	0.9494	0.9459	0.9467
		(150.2)	(130.7)	(110.5)
Weights	γ_2		0.0966	
			(2.82)	
	γ_4		-0.0215	
			(-11.45)	
DoF	ν			5.662
				(11.29)
AIC		1.7910	1.7514	1.7433

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 N-ME^{+} distributions, for the BP/$ exchange-rate daily returns. γ_s denotes the weighting parameter of the $s-th$ order polynomial in the ME distribution. DoF denotes degrees of freedom, and AIC is the mean of the AICs of the R estimations through the out-of-sample period. t Statistics calculated from robust standard errors are in parentheses below the parameter estimates.

FIGURE 1
Fitted unconditional distributions



Panel A



Panel B

Panel A shows the histogram and fitted distribution of the in-sample returns from Gaussian, Student's t and ME models. Panel B highlights the fit of the left tail.

We now proceed to compare the relative performance of the N-ME$^+$ model for forecasting h_t. The forecast accuracy is measured with respect to the out-of-sample squared residuals, $\{\widehat{u}_t^2\}_{t=R+1}^{T}$, by using the loss functions family proposed by Patton (2010).[7]This class of loss functions is shown to be robust to models ranking when using imperfect volatility proxies (as,

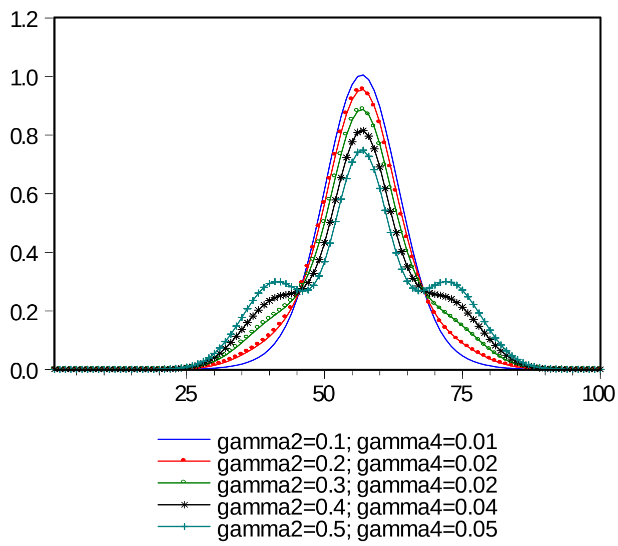
[7]Note that, if $R/(T-R-1) \to 0$, then the use of $\widehat{u}_t^2$ as the volatility proxy leads to the choice of the "right" model, as it does not alter the correct comparison of models, at least in terms of a quadratic loss function: see Awartani and Corradi (2005).

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.

$$L_{j,t}(\widehat{u}^2_{t+1}, \widehat{h}_{t+1}) = \begin{cases} e^2_{t+1}, & j = 1, \\ \widehat{u}^2_{t+1}/\widehat{h}_{t+1} - \log(\widehat{u}^2_{t+1}/\widehat{h}_{t+1}) - 1, & j = 2, \\ (\widehat{u}^6_{t+1} - \widehat{h}^3_{t+1})/6 - \widehat{h}^2_{t+1}(\widehat{u}^2_{t+1} - \widehat{h}_{t+1})/2, & j = 3. \end{cases} \quad (20)$$

where $e_{t+1} = \widehat{h}_{t+1} - \widehat{u}^2_{t+1}$.The significance of the difference between these loss functions is tested by using the Diebold and Mariano (DM) (1995) test.[8]

FIGURE 2
ME density allowable shapes



The Figure shows the allowable shape of the density in terms of the values of its parameters, gamma2 and gamma4 in the figure correspond to γ_2 and γ_4, respectively, in the text.

[8]For one 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 I and II is, $H_0 : E[p_{t+1}] = 0$, with $p_{t+1} = L_{j,t+1}(\widehat{u}_{t+1}, \widehat{h}^I_{t+1}) - L_{j,t+1}(\widehat{u}_{t+1}, \widehat{h}^{II}_{t+1})$. The test statistic is computed as: $DM = \overline{p}/(2\pi\widehat{\boldsymbol{\omega}}_p(\omega = 0)/R)^{\frac{1}{2}}$, where $\overline{p}$ is the sample mean of the loss differential series over the out-of-sample period, and $\widehat{\boldsymbol{\omega}}_p(\omega = 0)$ is a heteroscedasticity and autocorrelation robust estimator of the loss differential spectral density function at frequency 0. In the present context, as the prediction period R grows at a slower rate than the estimation period $T - R - 1$, $R/(T - R - 1) = 0.11$, then the effect of parameter estimation error vanishes and the DM statistic converges in distribution to a standard normal (see West, 1996).

Table 2 presents the results of the DM test for all pairwise comparisons and loss functions. The entries are the means of L_j over the out-of-sample period. The number in parenthesis below each entry is the p-value of the test. A sharp result that emerges from Table 2 is that there are not statistical differences between Gaussian and ME models but both significantly outperform the Student's t according to L_1 and L_3, respectively. A second observation is that the Student's t model tends to overpredict more and underpredicts less the volatility than the ME and the Gaussian models in this order, although differences are not statistically significant in relation to the error of overprediction loss function, L_2; these results are in line with those in the literature on volatility forecasting (see e.g. Ñíguez and Perote 2012).

TABLE 2
Out-of-sample volatility forecasting performance

Models	Gaussian	GME	Student's t
		L_1	
GME	0.15822 (0.200)		
Student's t		0.15847 (0.024)	
Gaussian			0.15905 (0.048)
		L_2	
GME	-0.41228 (1.207)		
Student's t		-0.41266 (0.312)	
Gaussian			-0.41141 (0.714)
		L_3	
GME	0.05586 (0.058)		
Student's t		0.05593 (0.007)	
Gaussian			0.05606 (0.012)

This table contains the results of the DM predictive ability test for the models and loss functions presented in this Section. 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.

Next, we test the models performance for forecasting r_t distribution tails. We compute R 1-step-ahead VaR forecasts for confidence levels $\alpha = \{0.1, 0.05, 0.025, 0.01\}$, $\widehat{VaR}^{\alpha}_{T+1} = \widehat{r}_{T+1} - \widehat{\xi}_T \widehat{h}^{1/2}_{T+1}$, where $\widehat{\xi}_T$ 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 for every in-sample window. The performance of the models is assessed by using the following criteria: unconditional coverage, $\widehat{\alpha}$, and the magnitude of the exception statistic, M_α, (López 1999), and the predictive quantile loss (PQL) function, Q_α; see, for instance, Giacomini and Komunjer (2005). 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 ME 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 latter model being generally preferred to the Student's t. A likely explanation for the good performance of the ME model may be the flexibility of the density expansions to parsimoniously fit the shape of the distribution tail. To illustrate these results, Figure 3 plots the VaR forecasts from the ME model against the out-of-sample returns.

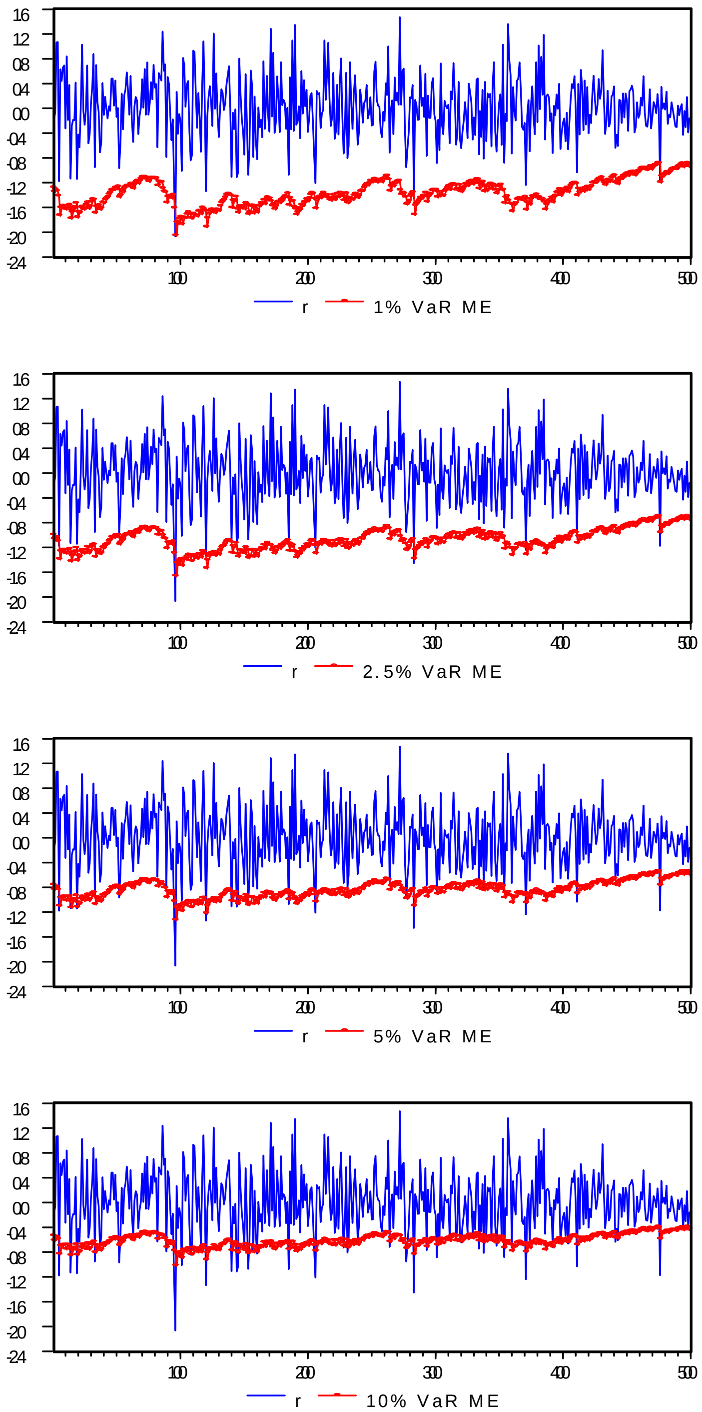
TABLE 3

VaR predictive accuracy

α	0.1	0.05	0.025	0.01
		Gaussian		
$\widehat{\alpha}$	0.1	0.046	0.02	0.006
C_α	7.1426	2.6904	1.0861	0.3537
Q_α	0.11632	0.07253	0.04344	0.02117
		N-ME$^+$		
$\widehat{\alpha}$	0.098	0.048	0.018	0.006
C_α	6.8856	2.7132	0.9735	0.2871
Q_α	0.11620	0.07304	0.04325	0.02087
		Student's t		
$\widehat{\alpha}$	0.102	0.048	0.024	0.006
C_α	7.1625	2.9044	1.1253	0.3160
Q_α	0.11685	0.07394	0.04423	0.02129

This table contains the results of the VaR tests. $\widehat{\alpha}$ denotes the estimated unconditional coverage probability, C_α denotes the magnitude of the exception statistic and, Q_α is the predictive quantile loss function, for one step ahead VaR forecasts with significance levels $\alpha = 0.1, 0.05, 0.025, 0.01$, obtained with the Gaussian, ME and Student's t models. Predictions 500.

FIGURE 3
VaR forecasts from N-ME^{+} model



Plots of 10%, 5%, 2.5% and 1% VaR forecasts from the N-ME^{+} model against out-of-sample BP/$ exchange-rate returns. Predictions 500.

4 Concluding remarks

Semi-nonparametric methods in general, and in particular Edgeworth and Gram-Charlier polynomial expansions, are typically applied to the Gaussian distribution. The resulting SNP pdf characterizes by its flexibility to parsimoniously capture small changes of frequency in distribution tails. However, SNP pdfs also present well-known theoretical and empirical rigidities, for instance, to expanding non-Normal pdfs (see Mauleón and Perote, 2000) or when ensuring positivity through, either Gallant and Nychka's type of transformations (Leon et al., 2009), or parametric constraints (Jondeau and Rockinger, 2001), because of the complexity and intractability of the resulting specifications.

In this paper, we propose a novel SNP method that is useful to address the aforementioned rigidities. We show that the resulting ME pdfs preserve the flexibility characteristic of GC expansions presenting besides advantageous features that pose them as an alternative to existing SNP pdfs. First, the ME has a very simple polynomial structure that does not require the orthogonality of its polynomials to prove its statistical properties. Second, the ME presents a general formulation that includes GC expansion as a special case. Third, the ME overcomes non-linearities among density parameters and moments when positive transformations are considered. Four, the ME naturally admits the use of non-normal distributions as basis with the only requirement of having as many finite moments as the expansion order.

We have analyzed the relative performance of the ME through an empirical application for forecasting the conditional variance and VaR of BP/$ exchange-rate returns, considering the Gaussian and Student's t pdfs as benchmark. The forecasts were evaluated by using the class of statistical loss functions in Patton (2010), the unconditional coverage and magnitude of the exceptions statistics in López (1999), and the PQL function in Koenker and Bassett (1978). Our results show that the ME model is as good as the Gaussian and both are significantly better than the Student's t for forecasting volatility. For forecasting VaR, we find evidence of that the ME model provides more accurate forecasts in relation to both the Student's t and the Engle's (2001) VaR-Gaussian models.

Appendix: ME properties

This Appendix presents the features of the ME in Propositions 1-9 and provides their proofs.

Proposition 1 A *ME of a pdf* $g(\cdot)$, denoted as $f(x,\boldsymbol{\gamma})$, *integrates to one:* $\int f(x,\boldsymbol{\gamma})dx = 1$.

Proof.

$$\int f(x,\boldsymbol{\gamma})dx = \int g\left(x\right)dx + \sum_{s=1}^{q}\gamma_s \int (x^s - \mu_s)g(x)dx = 1 + \sum_{s=1}^{q}\gamma_s(\mu_s - \mu_s) = 1. \tag{21}$$

■

Proposition 2 $0 \leq \gamma_s \leq \dfrac{1}{n\mu_s}$ $\forall s$ *even, and* $\gamma_s = 0$ $\forall s$ *odd, are sufficient conditions to guarantee the positiveness of* $f(x,\boldsymbol{\gamma})$.

Proof. The ME pdf defined in equation (6) can be re-written as,

$$f(x,\boldsymbol{\gamma}) = \left(\sum_{s=1}^{n}\gamma_s x^s + k\right)g(x), \tag{22}$$

where $k = 1 - \sum_{s=1}^{n}\gamma_s\mu_s$. Therefore, if $0 \leq \gamma_s \leq \frac{1}{n\mu_s}$ $\forall s$ even, and $\gamma_s = 0$ $\forall s$ odd, then $\sum_{s=1}^{n}\gamma_s x^s + k \geq 0$, since $\sum_{s=1}^{n}\gamma_s\mu_s \leq 1$. Consequently, $f(x,\boldsymbol{\gamma}) \geq 0$.■

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

$$m_i = E[x^i] = \mu_i + \sum_{s=1}^{n}\gamma_s(\mu_{s+i} - \mu_s\mu_i),\ \forall i = 1,2,... \tag{23}$$

Alternatively, the first n non-central moments of $f(x,\boldsymbol{\gamma})$ can be expressed in matrix form as,

$$\mathbf{M} = \boldsymbol{\mu} + \mathbf{A}\boldsymbol{\gamma}. \tag{24}$$

where $\mathbf{M} = (m_1, m_2, \ldots, m_n)'$ and $\mathbf{A}$ is a symmetric matrix of order n whose $i-th$ element is $\{a_{ij}\} = \{\mu_{s+i} - \mu_s\mu_i\}$.

Proof.

$$\begin{aligned} E[x^i] &= \int x^i f(x,\boldsymbol{\gamma})dx \qquad (25) \\ &= \int x^i g\left(x\right)dx + \sum_{s=1}^{n} \gamma_s \int x^i (x^s - \mu_s) g(x)dx = \mu_i + \sum_{s=1}^{n} \gamma_s (\mu_{s+i} - \mu_s\mu_i). \end{aligned}$$

■

Corollary 1 *The even/odd non-central moments of the ME of a symmetric pdf, $g(x)$,only depend on the even/odd parameters in γ. Therefore, the matrix A in equation (24) can be re-written as a block diagonal matrix, where the submatrices A_I and A_{II} contain the odd and even parameters of the ME pdf, $f(x,\gamma)$, respectively. Accordingly, the vectors M, μ and γ can also be partitioned in M_I and M_{II}, μ_I and μ_{II}, and γ_I and γ_{II}, respectively, containing, the even and the odd moments of $f(x,\gamma)$, the moments of $g(x)$ and the parameters of $f(x,\gamma)$, respectively. Equation (26) expresses equation (24) in terms of the resulting partitioned non-homogeneous equations system.*

$$\begin{bmatrix} \mathbf{M}_I \\ \mathbf{M}_{II} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_I \\ \boldsymbol{\mu}_{II} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{II} \end{bmatrix} \begin{bmatrix} \boldsymbol{\gamma}_I \\ \boldsymbol{\gamma}_{II} \end{bmatrix}. \tag{26}$$

Furthermore, if A is full-rank, then the system in equations (24) or (26) has the trivial *solution,*

$$\begin{aligned} \boldsymbol{\gamma} &= \mathbf{A}^{-1}(\mathbf{M}-\boldsymbol{\mu}), &(27)\\ \boldsymbol{\gamma}_I &= \mathbf{A}_I^{-1}(\mathbf{M}_I-\boldsymbol{\mu}_I) \text{ and } \boldsymbol{\gamma}_{II} = \mathbf{A}_{II}^{-1}(\mathbf{M}_{II}-\boldsymbol{\mu}_{II}). &(28)\end{aligned}$$

Corollary 2 *A ME of any parametric pdf, $f(x,\boldsymbol{\gamma})$, can be expressed in terms of its first n non-central moments as,*

$$f(x,\mathbf{M}) = \left[1+(\mathbf{Z}-\boldsymbol{\mu})'\mathbf{A}^{-1}(\mathbf{M}-\boldsymbol{\mu})\right]g(x) = \left(1+\sum_{s=1}^{n}(m_s-\mu_s)\theta_s(x)\right)g(x), \qquad (29)$$

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

Proposition 4 $\boldsymbol{\gamma} = \mathbf{A}^{-1}\mathbf{B}^{-1}\left(\mathbf{Sd}-(\mathbf{I}^*+\mathbf{B})\boldsymbol{\mu}^+\right)$ *is a necessary and sufficient condition for the Normal-ME pdf, $f_N(x,\boldsymbol{\gamma})$, and the GC density, $\pi(x,\mathbf{d})$, to have the same moments.*

Proof. The first n moments of the GC density can be obtained from equations (3) and (4) as follows:

$$\begin{aligned} \mathbf{M}_{GC} &= E[\mathbf{Z}] = \int \mathbf{Z}\left(1+\mathbf{H}'\mathbf{d}\right)\phi(x)dx &(30)\\ &= \int \mathbf{B}^{-1}(\mathbf{H}-\mathbf{I}^*\boldsymbol{\mu}^+)\left(1+\mathbf{H}'\mathbf{d}\right)\phi(x)dx\\ &= \mathbf{B}^{-1}\int \mathbf{H}\phi(x)dx\\ &\quad +\mathbf{B}^{-1}\int \mathbf{HH}'\mathbf{d}\phi(x)dx - \mathbf{B}^{-1}\mathbf{I}^*\boldsymbol{\mu}^+\int\phi(x)dx - \mathbf{B}^{-1}\mathbf{I}^*\boldsymbol{\mu}^+\int\mathbf{H}'\mathbf{d}\phi(x)dx\\ &= \mathbf{0}+\mathbf{B}^{-1}(\mathbf{S}\times\mathbf{d}) - \mathbf{B}^{-1}\mathbf{I}^*\boldsymbol{\mu}^+ + \mathbf{0}\\ &= \mathbf{B}^{-1}\left[(\mathbf{S}\times\mathbf{d})-\mathbf{I}^*\boldsymbol{\mu}^+\right].\end{aligned}$$

On the other hand, the first n moments of the Normal-ME, $f_N(x,\boldsymbol{\gamma})$, can be expressed as,

$$\mathbf{M}_{GME} = \mathbf{A}\boldsymbol{\gamma}+\boldsymbol{\mu}^+. \qquad (31)$$

Therefore, $\mathbf{M}_{ME} = \mathbf{M}_{GC}$ if and only if $\mathbf{A}\boldsymbol{\gamma} + \boldsymbol{\mu}^{+} = \mathbf{B}^{-1}\left[(\mathbf{S}\times\mathbf{d}) - \mathbf{I}^{*}\boldsymbol{\mu}^{+}\right]$ if and only if $\boldsymbol{\gamma} = \mathbf{A}^{-1}\mathbf{B}^{-1}\left(\mathbf{Sd} - (\mathbf{I}^{*}+\mathbf{B})\boldsymbol{\mu}^{+}\right)$.■

Proposition 5 *If $m_i^* = E[(x-m_1)^i]$ $\forall i = 1, 2, \ldots,$ are the central moments of $f(x,\boldsymbol{\gamma})$, then a standardized ME pdf, i.e. zero mean and variance one, denoted as $f(z,\cdot)$, can be defined either in terms of the density parameters, equation (32), or in terms of the density moments, equation (33).*

$$f(z,\boldsymbol{\gamma}) = \left(1 + \sum_{s=1}^{n} \gamma_s \Psi_s\left(m_2^{*1/2} z + m_1\right)\right) g\left(m_2^{*1/2} z + m_1\right) m_2^{*1/2}, \tag{32}$$

$$f(z,\mathbf{M}) = \left(1 + \sum_{s=1}^{n} (m_s - \mu_s)\theta_s\left(m_2^{*1/2} z + m_1\right)\right) g\left(m_2^{*1/2} z + m_1\right) m_2^{*1/2}. \tag{33}$$

Proof. If $x \sim f(x,\cdot)$, equation (6) or (29), then the standardized variable $z = \dfrac{x-m_1}{m_2^{*\frac{1}{2}}} \sim f^*(z,\cdot)$, equations (32) or (33), respectively.■

Proposition 6 *A positive ME of a given pdf $g(\cdot)$, $F(x,\boldsymbol{\gamma})$, can be obtained by squaring the polynomials of $f(x,\boldsymbol{\gamma})$ as,*

$$F(x,\boldsymbol{\gamma}) = \frac{1}{W}\left(1 + \sum_{s=1}^{n} \gamma_s^2 \Psi_s(x)^2\right) g(x), \tag{34}$$

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

$$W = \int \left(1 + \sum_{s=1}^{n} \gamma_s^2 \Psi_s(x)^2\right) g(x)dx = 1 + \sum_{s=1}^{n} \gamma_s^2(\mu_{2s} - \mu_s^2). \tag{35}$$

Proof.

$$\begin{aligned} W &= \int \left(1 + \sum_{s=1}^{n} \gamma_s^2 (x^s - \mu_s)^2\right) g(x)dx \qquad (36)\\ &= \int g(x)dx + \sum_{s=1}^{n} \gamma_s^2 \left(\int x^{2s} g(x)dx + \mu_s^2 \int g(x)dx - 2\mu_s \int x^s g(x)dx\right)\\ &= 1 + \sum_{s=1}^{n} \gamma_s^2(\mu_{2s} + \mu_s^2 - 2\mu_s^2) = 1 + \sum_{s=1}^{n} \gamma_s^2(\mu_{2s} - \mu_s^2). \blacksquare \end{aligned}$$

Proposition 7 *The non-central moments of $F(x,\boldsymbol{\gamma})$, denoted as $\widetilde{m}_i$, can be expressed in terms of the moments of the expanded density, $g(x)$, as displayed in equation (37).*

$$\widetilde{m}_i = E[x^i] = \mu_i + \sum_{s=1}^{n} \gamma_s^2 \left[\mu_{2s+i} + \mu_s(\mu_s\mu_i - 2\mu_{s+i})\right], \ \forall i = 1, 2, ... \tag{37}$$

Proof.

$$\begin{aligned} E\left[x^i\right] &= \int x^i \left(1 + \sum_{s=1}^{n} \gamma_s^2 (x^s - \mu_s)^2\right) g(x)dx \qquad (38) \\ &= \int x^i g(x)dx \\ &\quad + \sum_{s=1}^{n} \gamma_s^2 \left(\int x^{2s+i} g(x)dx + \mu_s^2 \int x^i g(x)dx - 2\mu_s \int x^{s+i} g(x)dx\right) \\ &= \mu_i + \sum_{s=1}^{n} \gamma_s^2 (\mu_{2s+i} + \mu_s^2\mu_i - 2\mu_s\mu_{s+i}) \\ &= \mu_i + \sum_{s=1}^{n} \gamma_s^2 \left[\mu_{2s+i} + \mu_s(\mu_s\mu_i - 2\mu_{s+i})\right]. \end{aligned}$$

■

Remark 1 Corollaries 1 and 2 apply to positive ME. It is noteworthy, however, that the positive GC pdf in León et al. (2005, 2009) and the Positive Edgeworth-Sargan pdf in Ñíguez and Perote (2012) cannot be reproduced from the positive ME with the same type of transformation. This is an important difference between GC and ME and one of the advantages of the latter expansion, which preserves nice properties (e.g. Proposition 7) even when positive transformations are implemented.

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

$$\begin{aligned} \Pr\left[x \leq a\right] &= \int_{-\infty}^{a} f_N(x,\boldsymbol{\gamma})dx = \int_{-\infty}^{a} \phi(x)dx \\ &\quad - \sum_{s=1}^{n} \gamma_s \left(a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \xi a^b\right) \phi(a), \end{aligned} \tag{39}$$

where

$$\xi = \begin{cases} (s-1)(s-3)\cdots 2, & \forall s \text{ odd}, \\ (s-1)(s-3)\cdots 3, & \text{otherwise}, \end{cases} \tag{40}$$

and

$$b = \begin{cases} 1, & \forall s \text{ even}, \\ 0, & \text{otherwise}. \end{cases} \tag{41}$$

Proof.

$$\begin{aligned}
\int_{-\infty}^{a} f_N(x,\boldsymbol{\gamma})dx &= \int_{-\infty}^{a} \left(1 + \sum_{s=1}^{n} \gamma_s (x^s - \mu_s^+)\right) \phi(x)dx \qquad (42)\\
&= \int_{-\infty}^{a} \left(1 - \sum_{s=1}^{n} \gamma_s \mu_s^+ + \sum_{s=1}^{n} \gamma_s x^s\right) \phi(x)dx \\
&= \int_{-\infty}^{a} \phi(x)dx \\
&- \sum_{s=1}^{n} \gamma_s \mu_s^+ \int_{-\infty}^{a} \phi(x)dx + \sum_{\substack{s=1 \\ s \text{ odd}}}^{n} \gamma_s \int_{-\infty}^{a} x^s \phi(x)dx + \sum_{\substack{s=1 \\ s \text{ even}}}^{n} \gamma_s \int_{-\infty}^{a} x^s \phi(x)dx \\
&= \int_{-\infty}^{a} \phi(x)dx \\
&- \sum_{s=1}^{n} \gamma_s \left[\left(x^{s-1} + (s-1)x^{s-3} + (s-1)(s-5)x^{s-5} + \ldots + \xi x^b\right) \phi(x)|_{-\infty}^{a} \right] \\
&= \int_{-\infty}^{a} \phi(x)dx - \sum_{s=1}^{n} \gamma_s \left(a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \xi a^b\right) \phi(a),
\end{aligned}$$

where,

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

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

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

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,

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

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

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

■

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

$$\begin{aligned} \Pr\left[x \le a\right] = \int_{-\infty}^{a} F_N(x,\boldsymbol{\gamma})dx &= \int_{-\infty}^{a} \phi(x)dx \\ &+ \frac{2}{W}\sum_{s=1}^{n} \gamma_s^2 \mu_s^+ \left(a^{s-1} + (s-1)a^{s-3} + (s-1)(s-3)a^{s-5} + \ldots + \zeta a^b\right)\phi(a) \\ &- \frac{1}{W}\sum_{s=1}^{n} \gamma_s^2 \left(a^{2s-1} + (2s-1)a^{s-3} + (2s-1)(2s-3)a^{2s-5} + \ldots + \mu_{2s}^+ a^b\right)\phi(a), \end{aligned} \tag{43}$$

where W is the constant in equation (35) for the moments of $\phi(x)$, denoted as μ_s^+, b is the constant defined in equation (41) and,

$$\zeta = \begin{cases} (s-1)(s-3)\cdots 2, & \forall s \textit{ odd}, \\ (s-1)(s-3)\cdots 3a, & \textit{otherwise}. \end{cases} \tag{44}$$

Proof.

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

where,

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

■

References

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

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

[3] Bollerslev, T. (1987): 'A conditional heteroskedastic time series model for speculative prices and rates of return.' *Review of Economics and Statistics,* 69, 542-547.

[4] Bollerslev, T., and J. M. Wooldridge (1992): 'Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances.' *Econometric Reviews,* 11, 143-172.

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

[6] Corrado, C., and T. Su (1997): 'Implied volatility skews and stock return Skewness and kurtosis implied by stock option prices.' *European Journal of Finance,* 3, 73-85.4.

[7] Del Brio, E. B., T. M. Ñíguez, and J. Perote (2009): 'Gram-Charlier Densities: A Multivariate Approach.' *Quantitative Finance,* 9, 855-868.

[8] Del Brio, E. B., T. M. Ñíguez, and J. Perote (2011): 'Multivariate semi-nonparametric densities with dynamic conditional correlations.' *International Journal of Forecasting,* 27, 347-364.

[9] Diebold, F. X., and R. Mariano (1995): 'Comparing predictive accuracy.' *Journal of Business and Economics Statistics,* 13, 253-264.

[10] Edgeworth, F. Y. (1896): 'The asymmetrical probability curve.' *Philosophical Magazine,* 41.

[11] Edgeworth, F. Y. (1907): 'On the representation of statistical frequency by series.' *Journal of the Royal Statistical Society,* series A, 80.

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

[13] Gallant, A. R., and D.W. Nychka (1987): 'Seminonparametric maximum likelihood estimation.' *Econometrica,* 55, 363-390.

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

[15] Giacomini, R., and I. Komunjer (2005): 'Evaluation and combination of conditional quantile forecasts.' *Journal of Business and Economic Statistics*, 23, 416-431.

[16] Harvey, C. R., and A. Siddique (1999): 'Autoregressive conditional skewness.' *Journal of Financial and Quantitative Analysis,* 34, 465-487.

[17] Jarrow, R., and A. Rudd (1982): 'Approximate option valuation for arbitrary stochastic processes.' *Journal of Financial Economics,* 10, 347-369.

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

[19] Kendall, M., and A. Stuart (1977): '*The advanced theory of statistics.*' Griffin & Co, London.

[20] Koenker, R., and G. Bassett (1978): 'Regression quantiles.' *Econometrica,* 46, 33-50.

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

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

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

[24] Mauleón, I. and J. Perote (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.

[25] Ñíguez, T. M., I. Paya, D. Peel, and J. Perote (2012): 'On the stability of the CRRA utility under high degrees of uncertainty.' *Economics Letters,* 115, 244-248.

[26] Ñíguez, T. M., and J. Perote (2012): 'Forecasting heavy-tailed densities with positive Edgeworth and Gram-Charlier expansions.' *Oxford Bulletin of Economics and Statistics,* 74, 600-627.

[27] Perote, J. (2004): 'The multivariate Edgeworth-Sargan density.' *Spanish Economic Review,* 6, 77-96.

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

[29] Sargan, J. D. (1975): 'Gram-Charlier approximations applied to t ratios of k-class estimators.' *Econometrica,* 43, 327-346.

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

[31] Velasco, C., and P. M. Robinson (2001): 'Edgeworth expansions for spectral density estimates and studentized simple mean.' *Econometric Theory,* 17, 497-539.

[32] Verhoeven, P., and M. McAleer (2004): 'Fat tails in financial volatility models.' *Mathematics and Computers in Simulation,* 64, 351-362.

[33] Vilhelmsson, A. (2009): 'Value at risk with time-varying variance, skewness and kurtosis –The NIG-ACD mode.' *Econometrics Journal,* 12, 82–104.

[34] West, K. (1996): 'Asymptotic inference about predictive ability.' *Econometrica,* 64, 1067-1084.