Gram-Charlier densities: A multivariate approach

Brio, Esther B. del; Niguez, Trino-Manuel; Perote, Javier

GRAM-CHARLIER DENSITIES: A MULTIVARIATE APPROACH

Journal:	*Quantitative Finance*
Manuscript ID:	RQUF-2007-0057.R1
Manuscript Category:	Research Paper
Date Submitted by the Author:	09-Jul-2008
Complete List of Authors:	DEL BRIO, ESTHER; UNIVERSITY OF SALAMANCA, Dept. of Business and Finance TRINO-MANUEL, ÑÍGUEZ; UNIVERSITY OF WESTMINSTER, Dept. of Economics and Quantitative Methods PEROTE, JAVIER; REY JUAN CARLOS UNIVERSITY, Dept. of Foundations of Economic Analysis
Keywords:	Financial Econometrics, Non-Gaussian Distributions, GARCH models, Forecasting Ability, Risk Management, Asymmetry
JEL Code:	C16 - Specific Distributions < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, C14 - Semiparametric and Nonparametric Methods < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, G1 - General Financial Markets < G - Financial Economics

Gram-Charlier Densities: A Multivariate Approach*

Esther B. Del Brio†
Department of Business and Finance, University of Salamanca, Salamanca, Spain

Trino-Manuel Ñíguez‡
Department of Economics and Quantitative Methods, University of Westminster, London, UK

Javier Perote§
Department of Foundations of Economic Analysis, University Rey Juan Carlos, Madrid, Spain

*This research has been supported by the Spanish Ministry of Education under grant SEJ2006-06104/ECON. A version of this paper is published as Fundación Cajas de Ahorros (FUNCAS) Working Paper No. 381. We wish to thank participants at the 1st Workshop on Computational and Financial Econometrics in Geneva, Switzerland, April 2007, and two anonymous referees, for their comments. All remaining errors are ours.

†E-mail: ebrio@usal.es

‡E-mail: T.M.Niguez@wmin.ac.uk

§Correspondence to: Javier Perote, Department of Foundations of Economic Analysis, Faculty of Law and Social Sciences, University Rey Juan Carlos, Campus de Vicálvaro, P^o de los Artilleros s/n, 28032 Madrid, Spain. Tel: +34 91 4887684. E-mail: javier.perote@urjc.es

Abstract

This paper introduces a new family of multivariate distributions based on Gram-Charlier and Edgeworth expansions. This family encompasses many of the univariate semi-nonparametric densities proposed in financial econometrics as marginal of its different formulations. Within this family, we focus on the analysis of the specifications that guarantee positivity to obtain well-defined multivariate semi-nonparametric densities. We compare two different multivariate distributions of the family with the multivariate Edgeworth-Sargan, Normal, Student's t and skewed Student's t in an in- and out-sample framework for financial returns data. Our results show that the proposed specifications provide a quite reasonably good performance being so of interest for applications involving the modelling and forecasting of heavy-tailed distributions.

Key words: Financial assets returns; Gram-Charlier and Edgeworth-Sargan densities; Leptokurtic multivariate distributions; MGARCH models; Skewness.

JEL classification: C16, G1.

1 Introduction

There is abundant literature on the non-normality of asset returns and its implications for pricing and measuring financial risk. Currently, decisions on capital allocation and portfolio management rely on computations of value-at-risk (VaR hereafter) measures or short-fall probabilities, for what normality or non-normality is a key assumption. Since Maldelbrot (1963), financial econometricians have analyzed the models misspecification related to the normality assumption, since it is certainly possible that models can produce the most accurate forecasts if the correct density is specified; recent developments on this line of research are provided in, for instance, Jurczenko et al. (2004), Jondeau and Rockinger (2005, 2006a, 2006b, 2007) and Boudt et al. (2007). These articles provide evidence on the importance of correctly accounting for, not only the time-varying dependence of conditional moments (i.e. conditional heteroskedasticity, skewness, or kurtosis), but also the shape of the whole underlying leptokurtic and possibly skewed density, specially that of the tails. Furthermore, they also highlight the convenience of modelling the joint portfolio distribution by assuming multivariate specifications and incorporating cross-moments structures (e.g. covariance, co-skewness, co-kurtosis).

Multivariate GARCH-type processes (MGARCH hereafter) have undergone important extensions since the constant conditional correlation (CCC) model of Bollerslev (1990) to the dynamic conditional correlation (DCC) model of Engle (2002) and Engle and Sheppard (2001); see Bauwens et al. (2005) for a complete survey on MGARCH models. On the other hand, different multivariate distributions have been introduced in financial econometrics, from which parametric approaches include: Student's t (Harvey et al., 1992), mixtures of Normals (Vlaar and Palm, 1993), skewed Normal (Azzalini and Dalla Valle, 1996), skewed Student's t (Sahu et al. (2003) and Bauwens and Laurent (2005)), Edgeworth-Sargan (Perote, 2004), Weibull (Malevergne and Sornette, 2004), Kotz-type (Olcay, 2005) and Normal Inverse Gaussian (Aas et al., 2006). Alternatively, any true target distribution can be approximated (fitted) through an infinite (finite) Gram-Charlier (GC hereafter) or Edgeworth series in terms of its moments or cumulants (see Sargan (1975, 1976) for the first applications of these techniques to econometrics). This semi-nonparametric (SNP hereafter) approach has the advantage of its general and flexible structure, since endogenously admits as much parameters as necessary depending on the empirical features of the data. Nonetheless, the applications of these distributions in finance, mainly for asset or option pricing, usually have not considered expansions beyond the fourth order (see, e.g., Corrado and Su (1996, 1997), Harvey and Siddique (1999), Jondeau and Rockinger (2001) and León et al. (2005)).[1] It is worth mentioning that, the application of SNP densities requires of the use of methods to ensure that the resulting truncated density is well-defined, i.e., it is positive for all values of its parameters in the parametric space. For this purpose different alternatives have been proposed in the literature depending on the end-use of the model, namely: *i)* accurate selection of initial values for the maximum likelihood algorithms (Mauleón and Perote, 2000), *ii)* parametric constraints (Jondeau and Rockinger, 2001), and *iii)* density function transformations based on the methodology of Gallant and Nychka (1987) and Gallant and Tauchen (1989). The first method is appropriate for in-sample analysis, whilst the

[1] On the other hand, Mauleón and Perote (2000) and Ñíguez and Perote (2004) have shown that expansions to the eighth order may provide a better goodness-of-fit.

second and the third are also useful for out-of-sample analysis (see, León et al. (2005) and Ñíguez and Perote (2004) for applications of the latter method in in- and out-of-sample contexts). However, it is known that parametric constraints may lead to sub-optimization, the model losing out some of its flexibility, whilst reformulations may lead to theoretically less tractable specifications. Empirical works on SNP densities has shown their superior performance with respect to different specifications used in finance in the univariate framework but, to the knowledge of the authors, much less is known on their performance in the multivariate context. In particular, Perote (2004) generalized the Edgeworth-Sargan distribution (ES hereafter) to the n-dimensional case defining the Multivariate ES (MES hereafter), and provided evidence on its goodness-of-fit for financial returns data, despite the MES function not really being a well-defined probability density function because for some parameter values it might be negative.

In this article we tackle those issues by presenting a general family of multivariate densities based on GC expansions (MGC densities hereafter) that is well-defined and encompasses most of the univariate GC densities proposed in the literature as marginals. We focus on two particular specifications that generalize to the multivariate framework the SNP and the Positive ES (PES hereafter) densities in León et al. (2007) and Ñíguez and Perote (2004), respectively. The theoretical properties of these densities (e.g. marginal distributions, cumulative distribution functions (cdf hereafter), univariate moments and cross-moments) are straightforwardly derived, showing that these distributions might be potentially superior in terms of flexibility to other alternative formulations, and analytically and empirically more tractable. The in-sample performance of the MGC specifications to fit financial data is compared to the MES, the Multivariate Normal (MN hereafter) and the (skewed) Multivariate Student's t ((Sk)-MST hereafter) through an empirical application to stock returns. We provide evidence on that the MGC distributions capture more accurately the heavy tails of portfolio returns distributions than the MN or the MST. This result is also obtained when the comparison is undertaken among skewed specifications (particularly we compare the asymmetric versions of the MGC and MES distributions with the Sk-MST). An application of the MGC densities for full density forecasting, based on the methodology in Diebold et al. (1998, 1999) and Davidson and MacKinnon (1998), is also provided. We compared a particular specification of our family with the MN, given its widely use by practitioners through the popular software package RiskMetrics of J.P. Morgan (1996). We show that the MGC densities provide a reasonably good performance for forecasting the full density of the portfolio and clearly overcomes the MN model.

The remainder of the article is structured as follows. Section 2 deals with the definitions and properties of the MGC family of densities. Section 3 tests the in- and out-of-sample performance of the proposed densities through an empirical application to a portfolio of stocks indexes, and Section 4 presents the main conclusions and suggests possible lines for further research.

2 Multivariate Gram-Charlier densities

In this section we introduce the family of MGC distributions, which is based on the SNP density approach derived from the Edgeworth and GC series. This family encompasses most of the univariate distributions based on expansions of this type used in the literature to model high-frequency financial returns for risk management purposes.

The "standardized" MGC family of densities is defined in terms of the "standardized" MN density, $G(\cdot)$, (i.e. with zero mean and unitary variance for all its marginal densities, $g(\cdot)$, and correlation coefficients denoted by ρ_{ij} $\forall i, j = 1, \ldots, n$, $i \neq j$), and the so-called Hermite polynomials, $H_s(\cdot)$, as given in the definition below.[2]

Definition 1 *A random vector $\mathbf{X} = (x_1, x_2, \ldots, x_n)' \in \mathbb{R}^n$ belongs to the MGC family of distributions if it is distributed according to the following density function,*

$$F(\mathbf{X}) = \frac{1}{n+1}G(\mathbf{X}) + \frac{1}{n+1}\left\{\prod_{i=1}^{n} g\left(x_i\right)\right\}\left\{\sum_{i=1}^{n}\frac{1}{c_i}\mathbf{h}\left(x_i\right)'\mathbf{A}_i\mathbf{h}\left(x_i\right)\right\}, \tag{1}$$

where A_i is a matrix of order $(q+1)$,[3] *$\mathbf{h}\left(x_i\right) = [1, H_1(x_i), H_2(x_i), \cdots, H_q(x_i)]' \in \mathbb{R}^{q+1}$, $H_s(\cdot)$ stands for the $s-th$ order Hermite polynomial described in equation (2),*

$$H_s(x_i) = \begin{cases} \sum\limits_{i=0}^{s/2}(-1)^i x_i^{s-2i}\frac{s!}{2^i i!(s-2i)!} & \forall s \text{ is even} \\ \sum\limits_{i=0}^{(s-1)/2}(-1)^i x_i^{s-2i}\frac{s!}{2^i i!(s-2i)!} & \text{otherwise,} \end{cases} \tag{2}$$

and c_i is the constant such that,

$$c_i = \int \mathbf{h}\left(x_i\right)'\mathbf{A}_i\mathbf{h}\left(x_i\right) g\left(x_i\right) dx_i. \tag{3}$$

The MGC family of functions straightforwardly integrates up to one and represents density functions providing that $\mathbf{A}_i$ is a positive definite matrix for all $i = 1, \ldots, n$. Within this family two straightforward but interesting cases deserve special attention. The first one arises when $\mathbf{A}_i$ admits the following decomposition, $\mathbf{A}_i = \mathbf{d}_i\mathbf{d}_i'$, with $\mathbf{d}_i = (1, d_{i1}, \ldots, d_{iq})' \in \mathbb{R}^{q+1}$ containing the density parameters (weights) of the $i-th$ density dimension. For this particular case, a positive version of the MGC density, hereafter named as MGCI, can be defined as in equation (4) below,

$$F_I(\mathbf{X}) = \frac{1}{n+1}G(\mathbf{X}) + \frac{1}{n+1}\left\{\prod_{i=1}^{n} g\left(x_i\right)\right\}\left\{\sum_{i=1}^{n}\frac{1}{c_i}\left[1 + \sum_{s=1}^{q} d_{is}H_s\left(x_i\right)\right]^2\right\}. \tag{4}$$

For this density, and based on the well-known orthogonality properties given in equations (5), (6) and (7),[4]

$$\int H_s(x_i)H_j(x_i)g(x_i)dx_i = 0 \quad \forall s \neq j, \tag{5}$$

$$\int H_s(x_i)H_j(x_i)g(x_i)dx_i = s! \quad \forall s = j, \tag{6}$$

$$\int H_s(x_i)g(x_i)dx_i = 0 \quad \forall s, \tag{7}$$

[2]Note that although we define the "standardised" MGC densities in terms of Gaussian densities with unitary variance, the resulting distributions do not have unitary variance, since variances, as the rest of the density moments, depend on the whole set of density parameters.

[3]Note that without loss of generality we have considered that for all dimensions the Gram-Charlier (Type A) expansions are truncated at the same order q.

[4]See Kendall and Stuart (1977) for further details about Hermite polynomials properties.

it can be straightforwardly proved that:

(i) The constant that weights the squared sum of Hermite polynomials for every variable x_i (see Proof 1 in the Appendix) is,

$$c_i = \int g(x_i) \left[1 + \sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 dx_i = 1 + \sum_{s=1}^{q} d_{is}^2 s! \tag{8}$$

(ii) The density integrates up to one, (see Proof 2 in the Appendix).

(iii) The marginal density for variable x_i is a mixture of a univariate Normal and a univariate SNP density of the type recently analyzed in León et al. (2005, 2007), as shown in equation (9) (see Proof 3 in the Appendix).

$$f_I(x_i) = \frac{n}{n+1} g(x_i) + \frac{1}{(n+1)c_i} \left[1 + \sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 g(x_i). \tag{9}$$

As a result of the last property, the moments of the distribution can be obtained immediately in terms of the moments of the normal distribution and the SNP density (see Fenton and Gallant (1996) or León et al. (2007) for a complete description of the moments of the SNP density). This fact also permits to introduce dynamic structures for the conditional moments of the distribution as proposed by Harvey and Siddique (1999) and León et al. (2005). For example, equation (10) considers conditional skewness for every variable i, s_{it}, in the "standardized" MGCI expanded to the third term.

$$F_I(\mathbf{X}) = \frac{1}{n+1} G(\mathbf{X}) + \frac{1}{n+1} \left\{\prod_{i=1}^{n} g(x_i)\right\} \left\{\sum_{i=1}^{n} \frac{1}{1 + \frac{s_{it}^2}{6}} \left[1 + \frac{s_{it}}{6}\left(x_i^3 - 3x_i\right)\right]^2\right\}. \tag{10}$$

Moreover, the covariances of this model are $\frac{1}{n+1}$ times the covariances of the MN process, $G(\cdot)$, and the co-skewness and co-kurtosis matrices can be worked out from the corresponding co-skewness and co-kurtosis matrices of the MN and the moments of the univariate normal and SNP distributions. To clarify this assessment we include the definitions of the co-skewness and co-kurtosis and an example of both types of cross-moments for the zero mean MGCI distribution (see Proof 4 in the Appendix). Particularly, the co-skewness (Harvey and Siddique, 2000) and co-kurtosis (Dittmar, 2002) matrices are defined in equations (11) and (12), respectively, and two examples for the MGCI are shown in equations (13) and (14).

$$\begin{aligned}
M_3 &= E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' \otimes (\mathbf{X}-\boldsymbol{\mu})'\right] = \{s_{ijk}\}, \ \forall i,j,k = 1,\dots,n, && (11)\\
M_4 &= E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' \otimes (\mathbf{X}-\boldsymbol{\mu})' \otimes (\mathbf{X}-\boldsymbol{\mu})'\right] = \{\kappa_{ijkl}\}, \ \forall i,j,k,l = 1,\dots,n, && (12)\\
s_{112} &= \frac{1}{n+1}\left\{s_{112}^* + E_{GC}\left[x_1^2\right] E_N[x_2] + E_N\left[x_1^2\right] E_{GC}[x_2] + E_N\left[x_1^2\right] E_N[x_2]\right\}, && (13)\\
\kappa_{1112} &= \frac{1}{n+1}\left\{\kappa_{1112}^* + E_{GC}\left[x_1^3\right] E_N[x_2] + E_N\left[x_1^3\right] E_{GC}[x_2] + E_N\left[x_1^3\right] E_N[x_2]\right\}, && (14)
\end{aligned}$$

where $\otimes$ stands for the Kronecker product, $\boldsymbol{\mu}$ is the mean vector, s_{ijk}^* and κ_{ijkl}^* represent the co-skewness and co-kurtosis for the MN distribution $\forall i,j,k,l = 1,\dots,n$, and $E_N[\cdot]$ and

$E_{GC}[\cdot]$ denote the expected value with respect to the univariate normal and univariate GC distribution, respectively. These results emphasize another potential advantage of using this family of distributions: It is not only their parametric flexibility to potentially improve data fits and incorporate different time-varying patterns for any moment (e.g. for modelling conditional skewness), but also their analytical simplicity. In fact, despite their apparently complex structure, the MGC distributions are theoretically easily tractable and easy to estimate by using the estimates of their marginal GC distributions as starting values for the optimization algorithms. In the next section, we provide empirical evidence supporting these issues, by illustrating the great flexibility of these densities to come out with varied shapes. We show that the MGC densities may present heavier tails than other distributions usually employed in finance, such as, the Student's t or the normal, besides being capable of capturing multimodality, what makes them very useful to accurately forecast risk measures related with the tails of assets returns distributions.

Furthermore, the MGCI distribution overcomes the aforementioned non-positivity problem that may arise when estimating the MES density in Perote (2004), equation (15).[5]

$$F_{ES}(\mathbf{X}) = G(\mathbf{X}) + \left\{\prod_{i=1}^{n} g(x_i)\right\}\left\{\sum_{i=1}^{n}\sum_{s=1}^{q} d_{is}H_s(x_i)\right\}, \tag{15}$$

The second case that is noteworthy arises when $\mathbf{A}_i = diag(1, d_{i1}^2, \ldots, d_{iq}^2)\ \forall i, j = 1, \ldots, n$. For these Hermite polynomials weighting matrices, the resulting density is defined in equation (16) below, which we denote as MGCII.

$$F_{II}(\mathbf{X}) = \frac{1}{n+1}G(\mathbf{X}) + \frac{1}{n+1}\left\{\prod_{i=1}^{n} g(x_i)\right\}\left\{\sum_{i=1}^{n}\frac{1}{c_i}\left[1 + \sum_{s=1}^{q} d_{is}^2 H_s(x_i)^2\right]\right\}. \tag{16}$$

Obviously this density is a particular case of the former formulation but it may result more useful and parsimonious in different applications. For this density the scaling constants, $c_i\ \forall i = 1, \ldots, n$, are also those in equation (8), but its marginals, displayed in equation (17), are mixtures of a univariate Normal and the univariate PES defined in Ñíguez and Perote (2004).

$$f_{II}(x_i) = \frac{n}{n+1}g(x_i) + \frac{1}{(n+1)c_i}\left[1 + \sum_{s=1}^{q} d_{is}^2 H_s(x_i)^2\right]g(x_i). \tag{17}$$

Therefore the MGCII distribution moments can be obtained as a combination of those of the univariate Gaussian and PES. Particularly the $k-th$ order PES even moment can be expressed as given in equation (18),

$$m_{ik} = \frac{1}{c_i}E_N[x_i^k] + \frac{1}{c_i}\sum_{s=1}^{n}\sum_{j=0}^{k/2} j!s!\delta_j d_s^2, \quad \forall k \text{ even}, \tag{18}$$

where $E_{MN}[x_i^k]$ denotes the $k-th$ order moment of the Gaussian density and $\{\delta_j\}_{j=0}^{k/2}$ is the sequence of constants that makes $x_i^k = \sum_{j=0}^{k/2}\delta_j H_j(x)^2$ (see Ñíguez and Perote (2004) for the details

[5]Note that for the maximum likelihood estimates the MES must be necessarily positive and thus this density can be estimated in many applications by choosing accurate initial values, based on the estimates for its marginal densities that are distributed as the univariate ES in Mauleón and Perote (2000).

of the moments of this distribution). For the sake of clarity, Table 1 includes the first four moments of the "standardized" PES expanded to the fourth order compared to the ES and SNP counterparts.

[**Table 1 Here**]

Regarding the cross-moments all the comments stated for the MGCI apply for the MGCII as well. Furthermore, the MGCII cdf can be easily worked out as shown in equation (19) (see Proof 5 in the Appendix), and consequently, they can be used easily for risk management purposes, either for modelling and forecasting credit risk, portfolio VaR or short-fall probabilities. The multivariate cdf of the MGCI can be obtained analogously in terms of the cdf of the univariate N(0,1) and univariate SNP distributions, see León et al. (2007) for further details.

$$\begin{aligned} & \Pr\left[x_1 \leq a_1, \cdots, x_n \leq a_n\right] \\ = \; & \frac{1}{n+1} \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} G(\mathbf{X}) dx_1 \cdots dx_n + \prod_{j=1,\; j\neq i}^{n-1} \int_{-\infty}^{a_j} g\left(x_j\right) dx_j \\ & \times \sum_{i=1}^{n} \left[\int_{-\infty}^{a_i} g\left(x_i\right) dx_i - \frac{g(a_i)}{c_i} \sum_{s=1}^{q} d_{is}^2 \sum_{k=0}^{s-1} \frac{s!}{(s-k)!} H_{s-k}(a_i) H_{s-k-1}(a_i) \right]. \end{aligned} \tag{19}$$

The MGC densities straightforwardly admit the specification of GARCH-type processes (Engle (1982) and Bollerslev (1986)) to explain the dynamics of their conditional moments. Particularly, the conditional variances, k_{it}^2, are introduced by considering transformations of the type $\mathbf{u}_t = \mathbf{\Lambda}_t(\boldsymbol{\alpha}) \cdot \mathbf{X}$ where $\mathbf{\Lambda}_t(\boldsymbol{\alpha}) = diag(k_{1t}, k_{2t}, \cdots, k_{nt})$.[6] Specifically, we consider the following specification,[7]

$$\mathbf{r}_t \;=\; \boldsymbol{\mu}_t(\mathbf{\Phi}) + \mathbf{u}_t, \tag{20}$$

$$\mathbf{u}_t|\mathbf{\Omega}_{t-1} \;\sim\; MGC(\mathbf{0}, \mathbf{\Sigma}_t(\boldsymbol{\theta})). \tag{21}$$

In the next section we test the performance of the bivariate versions of the MGCI and MGCII in comparison with the previous but "non-positive" attempt to generalize GC densities to the multivariate framework, i.e. the MES, and the most widely used distributions in finance: the MN, implemented in the popular software package RiskMetrics (J.P. Morgan, 1996), and the MST, which is thick-tailed for low values of the degrees of freedom parameter, υ. In addition, we also include the comparison of the asymmetric versions of the MES and MGCI densities (hereafter Sk-MES and Sk-MGCI, respectively) with the skewed MST (hereafter Sk-MST) in Bauwens and Laurent (2005). The "standardized" cases of the n-dimensional MST and Sk-MST are defined in equations (22) and (23), respectively.[8]

[6] Note that this model is a special case of the CCC model of Bollerslev (1990), where the constant correlations coefficients are the ρ_{ij} included in the "standardised" MN of the MGC densities.

[7] It must be noted that the covariance matrix of $\mathbf{r}_t$ is $\mathbf{\Sigma}_t(\boldsymbol{\theta}) = \mathbf{\Lambda}_t(\boldsymbol{\alpha})\mathbf{\Omega}_t(\mathbf{d},\boldsymbol{\rho})\mathbf{\Lambda}_t(\boldsymbol{\alpha})$, where $\mathbf{\Omega}_t(\mathbf{d},\boldsymbol{\rho})$ is the covariance matrix of the MGC process and $\boldsymbol{\theta}' = (\boldsymbol{\alpha}, \mathbf{d}, \boldsymbol{\rho})$, with $\boldsymbol{\alpha} > \mathbf{0}$ being the vector containing the parameters of the GARCH-type process for the conditional variances of $\mathbf{r}_t$, $\mathbf{d}$ the vector of weights, and $\boldsymbol{\rho}$ the vector of correlation parameters.

[8] See Kotz and Nadarajah (2004) for a complete survey on the existing alternative specifications of the MST distributions.

$$F_{ST}(\mathbf{X}|\upsilon) = \frac{\Gamma\left(\frac{\upsilon+n}{2}\right)}{(\pi(\upsilon-2))^{\frac{n}{2}}\Gamma\left(\frac{\upsilon}{2}\right)}\left[1+\frac{\mathbf{X}'\mathbf{X}}{\upsilon-2}\right]^{-\frac{\upsilon+n}{2}}. \quad (22)$$

$$F_{SST}(\mathbf{X}|\upsilon,\boldsymbol{\lambda}) = \frac{\Gamma\left(\frac{\upsilon+n}{2}\right)}{(\pi(\upsilon-2))^{\frac{n}{2}}\Gamma\left(\frac{\upsilon}{2}\right)}\left(\prod_{i=1}^{n}\frac{2b_i}{\lambda_i+\frac{1}{\lambda_i}}\right)\left[1+\frac{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{\upsilon-2}\right]^{-\frac{\upsilon+n}{2}}, \quad (23)$$

$$\begin{aligned}
\varepsilon_i &= (b_i x_i + a_i)\lambda_i^{-I_t}, \\
I_t &= \begin{cases} 1 & \text{if } x_i \geq -\frac{a_i}{b_i} \\ -1 & \text{if } x_i < -\frac{a_i}{b_i} \end{cases}, \\
b_i &= \left[\left(\lambda_i^2 + \frac{1}{\lambda_i^2} - 1\right) - a_i^2\right]^{1/2}, \\
a_i &= \frac{\Gamma\left(\frac{\upsilon-1}{2}\right)\sqrt{\upsilon-2}}{\sqrt{\pi}\Gamma\left(\frac{\upsilon}{2}\right)}\left(\lambda_i - \frac{1}{\lambda_i}\right),
\end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) > \mathbf{0}$ is the vector of asymmetry parameters, so $F_{SST}(\mathbf{X}|\upsilon,\boldsymbol{\lambda})$ is skewed to the right (left) if $ln(\lambda_i) > 1$ (< 1) and $F_{SST}(\mathbf{X}|\upsilon,\boldsymbol{\lambda})$ reduces to $F_{ST}(\mathbf{X}|\upsilon)$ when $\boldsymbol{\lambda} = \mathbf{0}$, and υ is constrained to be larger than 2 for ensuring the existence of the covariance matrix.[9]

3 An empirical application to portfolio returns

The data used are daily returns of S&P500 and the Hang-Seng indices of the New York and Hong Kong Stock Exchange, respectively, $\mathbf{r}_t = (r_{1t}, r_{2t})$, over the period December 19, 1991 to December 19, 2006 for a total of $T = 3,913$ observations, obtained from Datastream. Plots and descriptive statistics of $\mathbf{r}_t$ are presented in Figure 1.

[**Figure 1 Here**]

Let the conditional distribution of $\mathbf{r}_t$, be either MN, MES, MGCI, MGCII or MST, with conditional mean and covariance matrix modelled according to equations (20) and (21). In particular, we use an AR(1) process (selected according to the Schwarz Bayesian Information Criterion (BIC)) to filter the small structure presented by the conditional mean of $\mathbf{r}_t$, and a GARCH(1,1) process to account for volatility clustering in the conditional variance of $\mathbf{r}_t$, as shown in equations (24) and (25),

$$\begin{aligned}
\mu_{it} &= \phi_{i0} + \phi_{i1} r_{it-1}, \qquad \forall i = 1, 2, & (24)\\
k_{it}^2 &= \alpha_{i0} + \alpha_{i1} u_{t-1}^2 + \alpha_{i2} k_{it-1}^2, & (25)
\end{aligned}$$

The estimation procedure is carried out in two steps using an in-sample window of $S = 3,512$ observations. Firstly, the AR(1) process is estimated by ordinary least squares and, secondly,

[9] For the particular bivariate-$F_{SST}(\mathbf{X}|\upsilon,\boldsymbol{\lambda})$ case, $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\varepsilon_{1t}^2 + \varepsilon_{2t}^2 - 2\rho\varepsilon_{1t}\varepsilon_{2t}) / \left(1-\rho^2\right).$

covariance matrix coefficients and density function 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 using Bollerslev and Wooldridge (1992) formula. The Hermite polynomials expansions were truncated at the $8th$ term according to accuracy criteria.[10] For the purpose of concentrating on the heavy tails of the distribution and considering the fact that the odd parameters were found not statistically jointly significant, in a first approach these odd parameters were constrained to zero.[11]

The likelihood function is maximized using the Newton-Raphson method. We observe that the estimation of the MGC models is not computationally very demanding providing that starting values are chosen properly. As MGC models are nested, a usual procedure to choose those values is to start with the estimation of simpler specifications and use those estimates as starting values for the estimation of more complex models. This is important given the high nonlinearity of the likelihood functions of MGC models. On the other hand, as it is known that estimated MGC densities for stock returns may present multiple local modes, it is important to ensure that the numerical maximization of the likelihood function do not yield a local optimum. For this purpose the optimization is monitored using different starting values to ensure that the obtained ML estimates are global optimums.[12]

Table 2 displays the estimates and their corresponding t-statistics (in parentheses) of the parameters of the considered symmetric models. A first observation is that both indexes present a very small linear dependency in the conditional mean: the estimated unconditional mean is higher for the Hang-Seng ($\widehat{\phi}_{20} > \widehat{\phi}_{10}$), and the AR(1) slope coefficient is significantly higher for the S&P500. All AR(1) coefficients are not significant at 5% level, but $\widehat{\phi}_{10}$, $\widehat{\phi}_{20}$, $\widehat{\phi}_{11}$ are at 10% level. Secondly, we observe the typical estimates of GARCH processes for financial returns; for both indexes the GARCH parameters estimates of all models reflect the existence of clustering and high persistence in volatility ($\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2}$ near but smaller than one), although the sum ($\widehat{\alpha}_{i1} + \widehat{\alpha}_{i2}$) is significantly lower for the MCGII model, in line with the results in Ñíguez and Perote (2004).

[**Table 2 Here**]

In relation to the estimated correlations, $\widehat{\rho}$, they are also significant and slightly higher for MGC and MES models. It must be noted though that the correlation coefficients and the conditional variance parameters of those specifications have to be interpreted carefully. For example, ρ in the MGC does not capture exactly the correlation among both variables, which explains the differences of this parameter estimate among the MGC densities and, the MN and MST models. Specifically,

[10] The BIC was employed to decide on the optimal lenght of the expansions. The truncation order is consistent to other papers that use expansions beyond the fourth Hermite polynomial; see, e.g., Mauleón and Perote (2000), Ñíguez and Perote (2004) or Perote (2004).

[11] The corresponding Likelihood ratio (LR) test accepted the null hypothesis $H_0 : d_{i1} = d_{i3} = d_{i5} = d_{i7} = 0$ for all the considered densities. Those LR test results are not displayed in the text for the sake of simplicity but are available from the authors upon request.

[12] Monitored optimization is also used in the out-of-sample application below. Specifically, we proceed using the same starting value for all windows, instead of using the usual optimum from the previous data window. Of course, this mechanism is computationally inefficient, i.e., more time consuming, but it is necessary to avoid getting trapped in successive local optimums.

for the analyzed MGC densities, the Pearson's correlation coefficient can be computed as $\rho/3\sigma_{1t}\sigma_{2t}$ where σ_{it} is the standard error of every variable obtained from the corresponding univariate GC marginal densities, which, e.g., for the MGCII is given by,

$$\sigma_{it} = k_{it}\left[2/3 + (1/3)\frac{1 + 10d_{i2}^2 + 216d_{i4}^2 + 9360d_{i6}^2 + 685440d_{i8}^2}{1 + 2d_{i2}^2 + 24d_{i4}^2 + 720d_{i6}^2 + 40320d_{i8}^2}\right]^{1/2}, \quad \forall i = 1, 2. \tag{26}$$

Even more, the stationarity conditions of the GARCH processes in MGC distributions are also slightly different from the usual ones as well.[13]

In relation to the parameter weights estimates in Table 2 ($\widehat{d}_{is}$, $i = 1$ (S&P500), $i = 2$ (Hang-Seng), $s = 2, 4, 6, 8$) (hereafter $i, j = 1, 2$) we observe that most of them are significant at reasonable confidence levels, only clearly $\widehat{d}_{12}$ arises as not significant for the MGCII model and $\widehat{d}_{22}$ for all SNP models. These estimates explain that the portfolio distribution is highly leptokurtic and that the MGC models are able to capture parsimoniously that tail shape. This result is confirmed by the estimated degrees of freedom of the MST model, $\widehat{v}$, which equals 7.1. An interesting observation is that, the coefficient $\widehat{d}_{i8}$ is clearly significant in most of the cases, reinforcing the fact that the densities need to be expanded at least up to the $8th$ polynomial to capture the probabilistic mass in the extreme range of the tails. Note that although the interpretation of the parameters of the MGC densities requires a complete study of the distribution moments, it is clear that d_{ij} is linked to higher moments (i.e. heavier tails) the bigger the $j-th$ subindex is.

For the purpose of comparing the accuracy of the different specifications and as a first orientative approach, Table 2 includes the log-likelihood value ($\ln L$) and the BIC, computed as $-\ln L + p\ln(S)/2$, where p stands for the number of parameters of the model. According to these criteria the densities based on Edgeworth and GC expansions outperform the most popular distributions in finance (MN and MST). This improvement in accuracy is due to that MGC models present higher flexibility than MST models since they count with more parameters to parsimoniously account for the target distribution shape. Among the densities based on Edgeworth and GC series, the MGCI seems to provide the best fit in the whole domain. Nevertheless, this result does not necessarily imply the best performance in the tails.

An illustration of the allowable shapes of the MGCII density in comparison with the MN for their different ranges and domains, including the multimodality feature, is provided in Figure 2. Left plots (Figures 2.A, 2.C and 2.E) correspond to the fitted MGCII and right plots (Figures 2.B, 2.C and 2.F) to the MN. Particularly, Figures 2.A and 2.B represent the whole domain of the functions, and the rest of the figures illustrate details of the distributions tails. It is noteworthy the fact that the MGCII is capable of capturing different jumps in the probabilistic mass (see Figure 2.C) whilst for the same range the MN density decreases smoothly (see Figure 2.D). Furthermore, the MGCII captures more accurately the leptokurtic density behaviour since it assigns positive probability to areas in the tails where the MN does not (see Figures 2.E and 2.F).

[**Figure 2 Here**]

[13] See See Ñíguez and Perote (2004) for an example of the GARCH(1,1) stationarity conditions for the particular case of the univariate PES density.

These findings can also be illustrated by depicting the marginal densities of every variable computed from the estimates of the multivariate distributions. Figure 3 includes the fitted marginal density for S&P500 under different specifications (MN, MST, MES, MGCI and MGCII) in comparison to the histogram of the data. Figure 3.A represents the densities for the whole domain whilst Figure 3.B includes only the distributions left tails. From these pictures it is clear that although the MES seems to capture more accurately the sharply peaked density behaviour, the MGCII outperforms the other specifications in the tails. Specifically, this distribution is clearly superior to less flexible distributions such as the MN and MST and other semi-nonparametric alternatives (MES or MGCI).[14] Therefore in the out-of-sample application below we analyze the performance of the MGCII as a representative well-behaved MGC distribution.

[**Figure 3 Here**]

The aforementioned comparisons are focused on the tail behaviour of symmetric density specifications. Nevertheless financial returns also seem to feature skewness. This evidence was specially found when conditional skewness processes were incorporated in the modelling of the returns distribution (Harvey and Siddique,1999 and 2000). In order to account also for this feature we estimated the Sk-MES and Sk-MGCI distributions, expanded up to $8th$ term but also including the odd Hermite polynomials. The corresponding results are presented in Table 3, which also displays the estimates for the Sk-MST for comparison purposes.[15] Furthermore, to provide more evidence on the conditional dynamics of the skewness of financial returns, we extended the time-varying skewness approach of Harvey and Siddique (1999) (see also León et al., 2005) to the multivariate framework by estimating the model in equation (10) with conditional skewness following a GARCH-type process (equation (27)),

$$s_{it} = \gamma_{i0} + \gamma_{i1} \left(\frac{u_{it-1}}{k_{it-1}} \right)^3 + \gamma_{i2} s_{it-1}, \tag{27}$$

where $-1 < \gamma_{i0} < 1$ represents the unconditional skewness, and $\gamma_{i1} \in \mathbb{R}$ and $\gamma_{i2} \in \mathbb{R}$ gather the relationship between current skewness, s_{it}, and past shocks to skewness, $(u_{it-1}/k_{it-1})^3$, and lagged skewness, s_{it-1}, respectively. The stationarity condition for the conditional skewness is that $\gamma_{i1} + \gamma_{i2} < 1$. The estimates of the corresponding density, that we denote CSk-MGCI are also displayed in Table 3. Firstly, we note that the estimates of the conditional variance processes, and the correlation coefficients and degrees of freedom are similar to those of the corresponding symmetric models in Table 2. A second observation is that for the S&P500 index the skewness coefficient, $\widehat{\lambda}_1$, in the Sk-MST model is not significantly different from 1, meaning that the S&P500 returns density is unconditionally symmetric, whilst for the Hang-Seng, $\widehat{\lambda}_2$ is significantly smaller than 1 at 10% level, so the Hang-Seng returns density is slightly unconditionally skewed. This result is confirmed by the individual significance tests of the coefficients $\widehat{\gamma}_{10}$ and $\widehat{\gamma}_{20}$ in the CSk-MGCI model, and the even weights coefficients in the Sk-MES and Sk-MGCI models. However,

[14] Note that, as pointed by Mauleón and Perote (2000), the degrees of freedom of the MST might be understated in an attempt to capture both the sharp peak and heavy tails with only this parameter. This fact explains the misspecified tail behaviour of the MST.

[15] The AR(1) coefficients of models in Table 3 are not presented in that table for symplicity, since they are the same than those in Table 2 because of the two-steps estimation procedure.

the LR test cannot reject the null hypothesis $H_0 : d_{13} = d_{15} = d_{17} = d_{23} = d_{25} = d_{27} = 0$ at any reasonable significance level for both Sk-MES and Sk-MGCI distributions. An explanation of these results is that although the marginal density of the Hang-Seng index returns is slightly skewed, the magnitude of its skewness coefficient is not large enough so that the joint null hypothesis of symmetry for the bivariate distribution is rejected. Turning to the conditional skewness of the bivariate MGCI density, the estimates of $\widehat{\gamma}_{i1}$ and $\widehat{\gamma}_{i2}$ are positive and significant, what shows that for both indexes, days with high skewness are followed by days with high skewness, and shocks to skewness are significant, although they are less relevant than its persistence. This dynamics are similar to those of the conditional variance of the returns indexes, i.e., there is skewness clustering similar to volatility clustering, however the skewness persistence is lower than that observed in volatility, as the sum $\widehat{\gamma}_{i1} + \widehat{\gamma}_{i2}$ is not as near one as the corresponding one of the coefficients in the conditional variance process; this result is in line with those of León et al. (2005) for exchange rates. The BIC statistics in Table 3 are just orientative since they are only comparable for the Sk-MES and Sk-MGCI models.

[**Table 3 Here**]

Finally, we test the performance of the MGC densities for forecasting the full density of the portfolio and compare the forecasts with those of a MN model by using the methodology in Diebold et al. (1998, 1999) and Davidson and MacKinnon (1998). The application of this methodology in a multivariate framework is based on cdfs, evaluated at the forecasted standardized AR(1) residuals, $\widehat{u}_{it+1} = (r_{it+1} - \widehat{\mu}_{it+1})/\widehat{k}^{-1}_{it+1}$, through the out-of-sample period ($N = 400$ observations). The resulting so-called probability integral transforms (PITs) sequences, labelled p_{it}, $p_{i|jt}$, $\forall i, j = 1, 2$ are *i.i.d.* $U(0, 1)$ under correct density specification,

$$\begin{aligned} p_{it} &= \int_{-\infty}^{\widehat{u}_{it+1}} f_{it+1}(u_{it+1}) du_{it+1}, \\ p_{i|jt} &= \int_{-\infty}^{\widehat{u}_{it+1}} f_{i|jt+1}(u_{it+1}) du_{it+1} = \frac{\int_{-\infty}^{\widehat{u}_{it+1}} \int_{-\infty}^{\widehat{u}_{jt+1}} f_{t+1}(u_{it+1}, u_{jt+1}) du_{it+1} du_{jt+1}}{\int_{-\infty}^{\widehat{u}_{jt+1}} f_{jt+1}(u_{jt+1}) du_{jt+1}}, \end{aligned} \tag{28}$$

where $f_{it}(\cdot)$, $f_{i|jt}(\cdot)$ and $f_t(\cdot)$ denote marginal, conditional and joint distributions, respectively. Moreover since p_{it} is also interpreted as the p-value corresponding to the quantile $\widehat{u}_{it+1}$ of the forecasted density we use the p-value plot methods in Davidson and MacKinnon (1998) to compare the models forecasting performance.[16] So, if the model is correctly specified the difference between the cdf of p_{it} and the 45^0 line should tend to zero asymptotically. The empirical distribution function of p_{it} can be easily computed as,

$$\widehat{P}_{p_{it}}(y_\varrho) = \frac{1}{N} \sum_{t=1}^{\mathbf{N}} \mathbf{1}(p_{it} \leq y_\varrho), \tag{29}$$

where $\mathbf{1}(p_{it} \leq y_\varrho)$ is an indicator function that takes the value 1 if its argument is true and 0 otherwise, and y_ϱ is an arbitrary grid of $\boldsymbol{\varrho}$ points, which it is made finer on its extremes to highlight

[16] Note that Davidson and MacKinnon (1998) used this method to compare the size and power of hypothesis tests, while following Fiorentini *et al.* (2003) we use it to discriminate among alternative models according to their performance for forecasting the full density.

models differences in the goodness-of-fit of the density tails. Alternatively, the p-value discrepancy plot (i.e. plotting $\widehat{P}_{p_{it}}(y_\varrho)-y_\varrho$ against y_ϱ) can be more revealing when it is necessary to discriminate among specifications that perform similarly in terms of the p-value plot (see Fiorentini et al., 2003). Consequently, under correct density specification, the variable $\widehat{P}_{p_{it}}(y_\varrho)-y_\varrho$ must converge to zero.

In Figures 4 and 5 we plot the marginal and conditional cdfs for the PIT series under either the MN (red line) or the MGCII (blue line) densities. A sharp observation that emerges from those graphs is that the MGCII model provides a reasonably good performance for forecasting the full density of the portfolio and clearly overcomes the MN model commonly used in financial risk applications.

[**Figures 4 and 5 Here**]

4 Concluding remarks

This paper introduces a family of multivariate distributions based on Edgeworth and GC expansions. This family encompasses most of the univariate densities proposed in financial literature (e.g. the so-called SNP or PES distributions), which can be obtained as the marginal densities of the different densities nested in this family. Therefore, the MGC densities inherit the properties of their univariate precursories in terms of their flexible parameter structure to accurately represent all the characteristic features of most high-frequency financial variables (i.e. thick tails, sharp peak, asymmetries, multimodality, conditional heteroskedasticity, etc.). The distributions of the family are necessarily positive since they can be understood as extensions of the Gallant and Nychka (1987) methodology to the multivariate framework. Therefore these formulations overcome the deficiencies of the MES density, which was the previous attempt to generalise the ES density to a multivariate framework.

The performance of these densities is compared to fit and forecast the full density of a portfolio of asset returns, and it is found that they perform quite satisfactorily and are superior to the MN and the MST (or skewed versions), the most commonly used distributions in financial risk management. Within the multivariate densities based on Edgeworth and GC expansions the MGCI seems to be more accurate than the other formulations. Moreover this specification allows the consideration of conditional time-varying skewness and thus the generalization of Harvey and Siddique (1999) model.

Nevertheless, the good performance in terms of accuracy measures in the whole domain do not necessarily imply the best fit in the distribution tails. We show that in some cases other more parsimonious specifications, such as MGCII, provide a better adjustment in the tails (although at the cost of a loss in accuracy when accounting for the skewness or the sharp peak in the mean). Therefore the choice among the different possibilities within the family depends not only on accuracy issues but also on other empirical and econometric considerations.

This paper opens a hopefully fruitful line of research providing general formulations for MGC densities, and showing evidence of their reasonably good in- and out-sample performance through an empirical application. These results suggest that although the MGC distributions could be an interesting tool for risk management further research seems worthwhile at both theoretical and empirical level, e.g. to improve data fits by considering dynamic structures for other moments (e.g. correlations or kurtosis) and to investigate the model performance for other financial applications, such as asset pricing or credit and market risk forecasting.

Appendix

This appendix includes the proofs of some properties of the MGC densities. Particularly, the constant that makes both the MGCI and the MGCII densities integrate up to one, the marginal densities and the cross-moments of the MGCI distribution and the cdf for the MGCII are derived. The corresponding proofs for other multivariate densities of the same family can be obtained likewise.

Proof 1: The constants that make MGCI and the MGCII integrate up to one are $c_i = 1 + \sum_{s=1}^{q} d_{is}^2 s!$, $\forall i = 1, \ldots, n$.

$$\begin{aligned}
c_i &= \int g(x_i)dx_i + \int \left[\sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 g(x_i)dx_i + 2\sum_{s=1}^{q} d_{is} \int H_s(x_i)g(x_i)dx_i \\
&= \int g(x_i)dx_i + \int \left[\sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 g(x_i)dx_i \\
&= 1 + \sum_{s=1}^{q}\sum_{j=1}^{q} d_{is}d_{ij} \int H_s(x_i)H_j(x_i)g(x_i)dx_i \\
&= 1 + \sum_{s=1}^{q} d_{is}^2 \int H_s(x_i)^2 g(x_i)dx_i = 1 + \sum_{s=1}^{q} d_{is}^2 s! \;\blacksquare
\end{aligned}$$

Proof 2: The MGCI density integrates up to one provided that c_i are the constants in Proof 1.

$$\begin{aligned}
\int \cdots \int F_I(\mathbf{X})dx_1 \cdots dx_n &= \frac{1}{n+1} \int \cdots \int G(\mathbf{X})dx_1 \cdots dx_n \\
&+ \frac{1}{n+1} \int \cdots \int \left\{\prod_{i=1}^{n} g\left(x_i\right)\right\} \left\{\sum_{i=1}^{n} \frac{1}{c_i} \left[1 + \sum_{s=1}^{q} d_{is} H_s\left(x_i\right)\right]^2\right\} dx_1 \cdots dx_n \\
&= \frac{1}{n+1} + \frac{1}{n+1} \sum_{i=1}^{n} g\left(x_i\right) \frac{1}{c_i} \left[1 + \sum_{s=1}^{q} d_{is} H_s\left(x_i\right)\right]^2 \\
&= \frac{1}{n+1} + \frac{n}{n+1} = 1 \;\blacksquare
\end{aligned}$$

Proof 3: The marginal densities of MGCI are convex combinations of a normal density and a univariate SNP density.

$$
\begin{aligned}
f_I(x_i) &= \int \cdots \int F_I(\mathbf{X}) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \\
&= \frac{1}{n+1} \int \cdots \int G(\mathbf{X}) dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \\
&+ \frac{1}{(n+1)c_i} g(x_i) \left[1 + \sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 \int \cdots \int \prod_{j=1,\ j\neq i}^{n} g(x_j)\, dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \\
&+ \frac{1}{n+1} g(x_i) \sum_{j=1,\ j\neq i}^{n} \frac{1}{c_j} \int \cdots \int \left[\prod_{h=1,\ h\neq i}^{n} g(x_h)\right] \left[1 + \sum_{s=1}^{q} d_{js} H_s(x_j)\right]^2 dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n \\
&= \frac{1}{n+1} g(x_i) + \frac{1}{n+1} \frac{1}{c_i} g(x_i) \left[1 + \sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 + \frac{n-1}{n+1} g(x_i) \\
&= \frac{n}{n+1} g(x_i) + \frac{1}{n+1} \frac{1}{c_i} g(x_i) \left[1 + \sum_{s=1}^{q} d_{is} H_s(x_i)\right]^2 \blacksquare
\end{aligned}
$$

Proof 4: The co-skewness of the MGCI density can be obtained in terms of the corresponding co-skewness of the MN density and the univariate moments of both the normal and SNP distribution.

$$
\begin{aligned}
s_{113} &= \int \cdots \int x_1^2 x_2 F_I(\mathbf{X}) dx_1 \cdots dx_n \\
&= \frac{1}{n+1} \int \cdots \int x_1^2 x_2 G(\mathbf{X}) dx_1 \cdots dx_n \\
&+ \frac{1}{(n+1)} \int x_1^2 \frac{1}{c_1} g(x_1) \left[1 + \sum_{s=2}^{q} d_{1s} H_s(x_1)\right]^2 dx_1 \int x_2 g(x_2)\, dx_2 \prod_{j=3}^{n} \left[\int g(x_j)\, dx_j\right] \\
&+ \frac{1}{(n+1)} \int x_1^2 g(x_1)\, dx_1 \int x_2 \frac{1}{c_2} g(x_2) \left[1 + \sum_{s=2}^{q} d_{2s} H_s(x_2)\right]^2 dx_1 \prod_{j=3}^{n} \left[\int g(x_j)\, dx_j\right] \\
&+ \frac{1}{(n+1)} \sum_{i=3}^{n} \left[\int x_1^2 g(x_1) dx_1 \int x_2 g(x_2) dx_2 \int \frac{1}{c_i} g(x_i) \left[1 + \sum_{s=2}^{q} d_{is} H_s(x_i)\right]^2 dx_i \left[\prod_{j=1, j\neq 1,2,i}^{n} \int g(x_j)\, dx_j\right]\right] \\
&= \frac{1}{n+1} \left\{s_{113}^* + E_{GC}\left[x_1^2\right] E_N\left[x_2\right] + E_N\left[x_1^2\right] E_{GC}\left[x_2\right] + E_N\left[x_1^2\right] E_N\left[x_2\right]\right\} \blacksquare
\end{aligned}
$$

Note that s_{ijk}^* stands for the co-skewness of the MN distribution $\forall i, j, k = 1, \ldots, n$ and $E_N[\cdot]$ and $E_{GC}[\cdot]$ denote the expected value with respect to the univariate normal and univariate GC distribution, respectively. The other cross-moments (e.g. $\kappa_{1112} = E_{MGC}\left[x_1^3 x_2\right]$ in equation (14)) are obtained likewise.

Proof 5: The cdf of the MGCII density can be obtained in terms of the cdf of the MN density and the cdf of the univariate normal and PES distribution.

$$\int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} F_{II}(\mathbf{X}) dx_1 \cdots dx_n = \frac{1}{n+1} \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} G(\mathbf{X}) dx_1 \cdots dx_n$$

$$+ \frac{1}{n+1} \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} \left[\prod_{i=1}^{n} g(x_i)\right] \left\{\sum_{i=1}^{n} \frac{1}{c_i}\left[1 + \sum_{s=1}^{q} d_{is}^2 H_s(x_i)^2\right]\right\} dx_1 \cdots dx_n$$

$$= \frac{1}{n+1} \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} G(\mathbf{X}) dx_1 \cdots dx_n$$

$$+ \frac{1}{n+1} \sum_{i=1}^{n} \left[\int_{-\infty}^{a_i} \frac{1}{c_i} g(x_i) \left[1 + \sum_{s=1}^{q} d_{is}^2 H_s(x_i)^2\right] dx_i \prod_{\substack{j=1 \\ j \neq i}}^{n} \int_{-\infty}^{a_j} g(x_j) dx_j\right]$$

$$= \frac{1}{n+1} \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} G(\mathbf{X}) dx_1 \cdots dx_n +$$

$$+ \frac{1}{n+1} \sum_{i=1}^{n} \left[\left[\int_{-\infty}^{a_i} g(x_i) dx_i - \frac{g(a_i)}{c_i} \sum_{s=1}^{q} d_{is}^2 \sum_{k=0}^{s-1} \frac{s!}{(s-k)!} H_{s-k}(a_i) H_{s-k-1}(a_i)\right] \prod_{j=1,\, j \neq i}^{n} \int_{-\infty}^{a_j} g(x_j) dx_j\right] \blacksquare$$

See Ñíguez and Perote (2004) for the details of the proof for the cdf of the PES density.

Tables

Table 1. Moments of the univariate ES, SNP and PES distributions defined in terms of $H_3(x_i) = x_i^3 - 3x_i$ and $H_4(x_i) = x_i^4 - 6x_i^2 + 3$.

	ES	SNP	PES
$E[x_i]$	0	$\frac{48d_{i3}d_{i4}}{1+6d_{i3}^2+24d_{i4}^2}$	0
$E[x_i^2]$	1	$\frac{1+42d_{i3}^2+216d_{i4}^2}{1+6d_{i3}^2+24d_{i4}^2}$	$\frac{1+42d_{i3}^2+216d_{i4}^2}{1+6d_{i3}^2+24d_{i4}^2}$
$E[x_i^3]$	$6d_{i3}$	$\frac{12d_{i3}^2+576d_{i3}d_{i4}}{1+6d_{i3}^2+24d_{i4}^2}$	0
$E[x_i^4]$	$3+24d_{i3}$	$\frac{3+450d_{i3}^2+2952d_{i4}^2+48d_{i4}}{1+6d_{i3}^2+24d_{i4}^2}$	$\frac{3+450d_{i3}^2+2952d_{i4}^2+48d_{i4}}{1+6d_{i3}^2+24d_{i4}^2}$

Table 2. Multivariate symmetric densities.

	MN	MST	MES	MGCI	MGCII
ϕ_{10}			.0341 (1.20)*		
ϕ_{11}			.0272 (1.65)*		
α_{10}	.0222 (2.49)	.0169 (2.39)	.0173 (1.90)	.0275 (2.59)	.0136 (2.49)
α_{11}	0722 (6.02)	.0542 (5.40)	.0551 (4.06)	.0712 (4.03)	.0422 (6.02)
α_{12}	.9209 (71.4)	.9390 (82.5)	.9299 (50.6)	.9277 (51.8)	.9208 (71.4)
ϕ_{20}			.0321 (1.89)*		
ϕ_{21}			-.0138 (-.81)*		
α_{20}	.0059 (2.20)	.0033 (2.08)	.0042 (2.26)	.0033 (2.49)	.0031 (2.20)
α_{21}	.0592 (4.88)	.0466 (5.14)	.0564 (5.26)	.0529 (5.68)	.0477 (4.88)
α_{22}	.9369 (74.4)	.9517 (102.4)	.9411 (89.3)	.9407 (102.3)	.9364 (74.4)
d_{12}			.2175 (1.13)	-.5019 (-7.13)	.5450 (1.08)*
d_{14}			.2602 (4.84)	.1581 (6.35)	.0350 (4.62)
d_{16}			.0458 (3.19)	-.0130 (-4.65)	-.0086 (2.94)
d_{18}			.0062 (4.42)	.0009 (3.76)	.0007 (4.17)
d_{22}			.0001 (0.71)*	-.0048 (-.10)*	.3149 (.58)*
d_{24}			.1370 (5.04)	.0360 (7.11)	.0001 (4.45)
d_{26}			.0187 (2.43)	-.0022 (-2.95)	-.0013 (1.33)*
d_{28}			.0035 (4.56)	.0005 (1.41)*	.0006 (3.10)
ρ	.1097 (6.22)	.1078 (6.49)	.2859 (6.61)	.2672 (6.58)	.3323 (6.22)
υ		7.01 (11.5)			
LnL	-10701.8	-10552.6	-10527.0	-10500.4	-10554.5
BIC	21432.3	21137.8	21115.2	21062.1	21170.2

Bivariate density for S&P500 (variable 1) and Hang-Seng (variable 2) indices. t-ratios in parentheses. The asterisk denotes approximate non-significance at 5% confidence level.

Table 3. Multivariate skewed densities.

	Sk-MST	Sk-MES	Sk-MGCI	CSk-MGCI
α_{10}	.0171 (2.39)	.0161 (1.98)	.0251 (2.52)	.0188 (2.68)
α_{11}	.0545 (5.40)	.0542 (3.78)	.0675 (4.17)	.0457 (5.11)
α_{12}	.9384 (82.1)	.9319 (49.2)	.9324 (57.6)	.9411 (87.6)
α_{20}	.0035 (2.13)	.0034 (1.97)	.0029 (2.33)	.0041 (1.98)
α_{21}	.0476 (5.20)	.0630 (4.95)	.0554 (6.15)	.0441 (2.15)
α_{22}	.9506 (101)	.9367 (83.7)	.9445 (105.1)	.9478 (47.1)
γ_{10}				-.0081 (-.78)*
γ_{11}				.0539 (5.20)
γ_{12}				.8792 (21.7)
γ_{20}				-.1285 (-2.66)
γ_{21}				.0952 (4.84)
γ_{22}				.6329 (2.95)
λ_1	.9941 (62.2)			
d_{12}		.2126 (1.08)*	-.5068 (-7.13)	
d_{13}		-.0298 (-.77)*	-.0096 (-.51)*	
d_{14}		.2579 (4.62)	.1575 (6.34)	
d_{15}		-.0208 (-1.27)*	-.0030 (-.78)*	
d_{16}		.0448 (2.94)	-.0132 (-4.66)	
d_{17}		-.0015 (-.67)*	-.0003 (-.95)*	
d_{18}		.0061 (4.17)	.0009 (3.51)	
λ_2	.9622 (42.4)			
d_{22}		.0001 (0.58)*	-.0711 (-2.94)	
d_{23}		-.1014 (-2.45)	-.0022 (-.25)*	
d_{24}		.1216 (4.45)	.0264 (4.88)	
d_{25}		-.0429 (-2.33)	.0012 (.59)*	
d_{26}		.0111 (1.33)*	-.0065 (-8.53)	
d_{27}		-.0052 (-2.16)	.0009 (3.83)	
d_{28}		.0002 (3.10)	.0001 (1.16)*	
ρ	.1059 (6.35)	.2881 (6.70)	.2717 (6.62)	.3187 (6.92)
υ	7.07 (11.5)			
LnL	-10550.8	-10520.6	-10497.4	-10623.3
BIC	21142.4	21126.9	21080.5	21299.6

Bivariate density for S&P500 (variable 1) and Hang-Seng (variable 2) indices. t-ratios in parentheses. The asterisk denotes approximate non-significance at 5% confidence level.

Figures

Figure 1. S&P 500 and Hang-Seng indices daily returns. Sample: 19/12/1991 - 06/06/2005 (observations 3,513). Out-of-sample: 07/06/2005 - 19/12/2006 (observations 400).

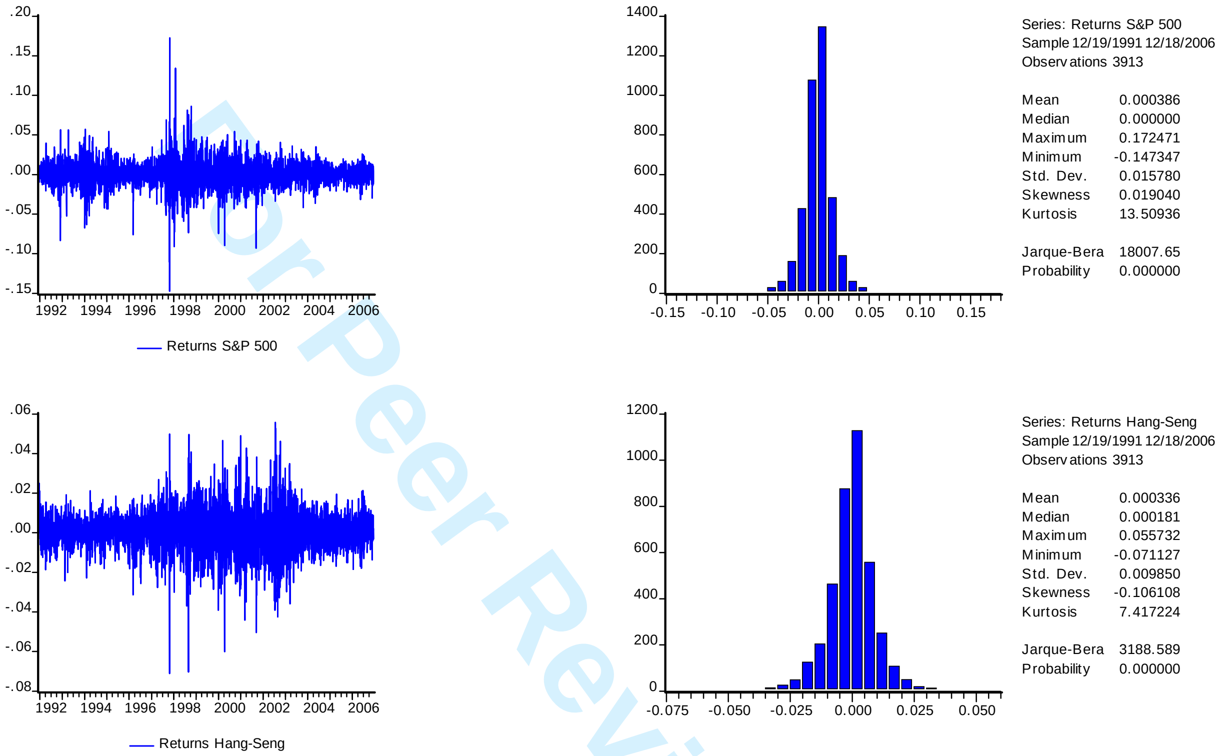


Figure 2: Fitted bivariate MN and MGCII densities of the Hang-Seng and S&P500 indexes returns for different ranges and domains.

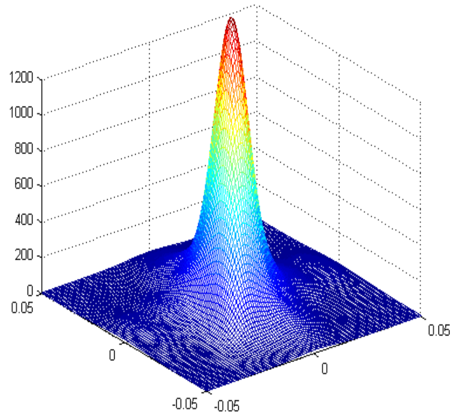


Figure 2.A

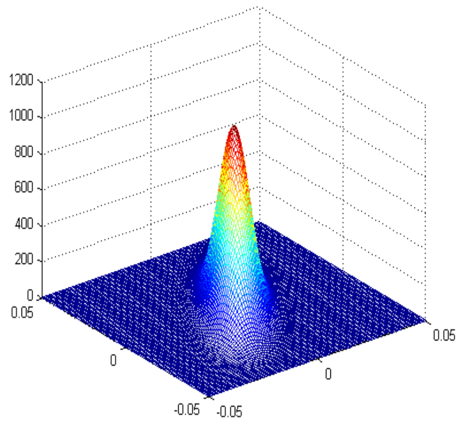


Figure 2.B

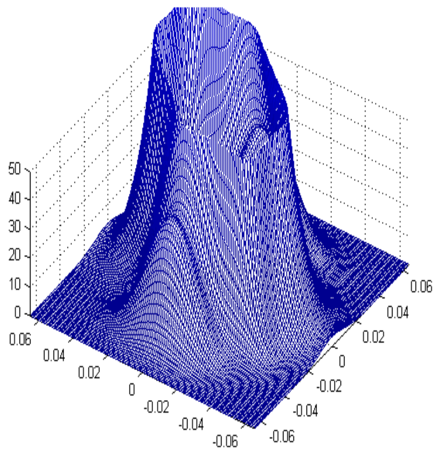


Figure 2.C

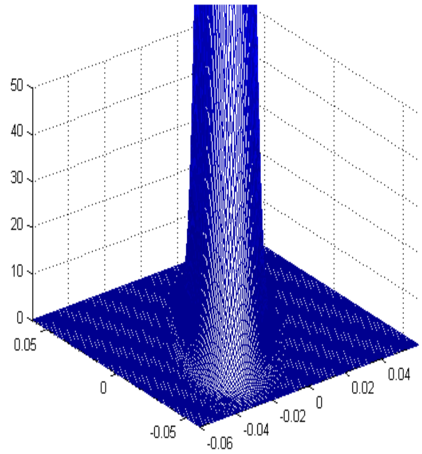


Figure 2.D

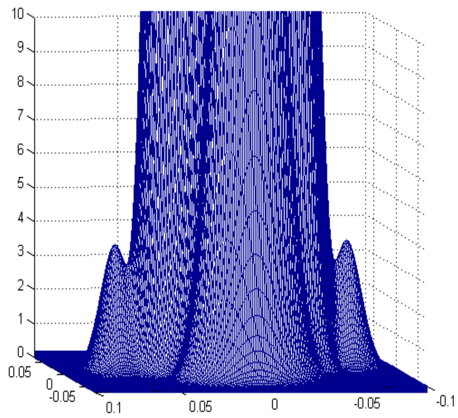


Figure 2.E

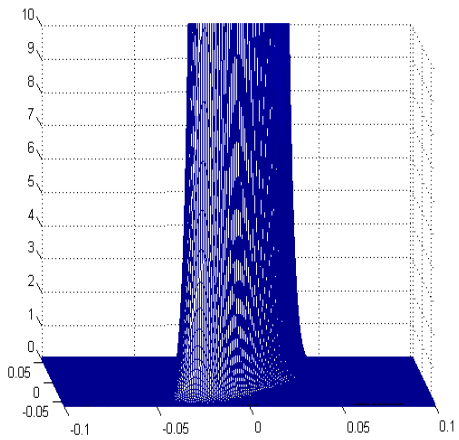


Figure 2.F

The first column of the figure corresponds to the fitted bivariate-MGCII density, and the second column to the fitted bivariate-MN density.

Figure 3: Fitted marginal density of S&P500 index computed from the estimated MN, MST, MES, MGCI and MGCII compared to the histogram of the data.

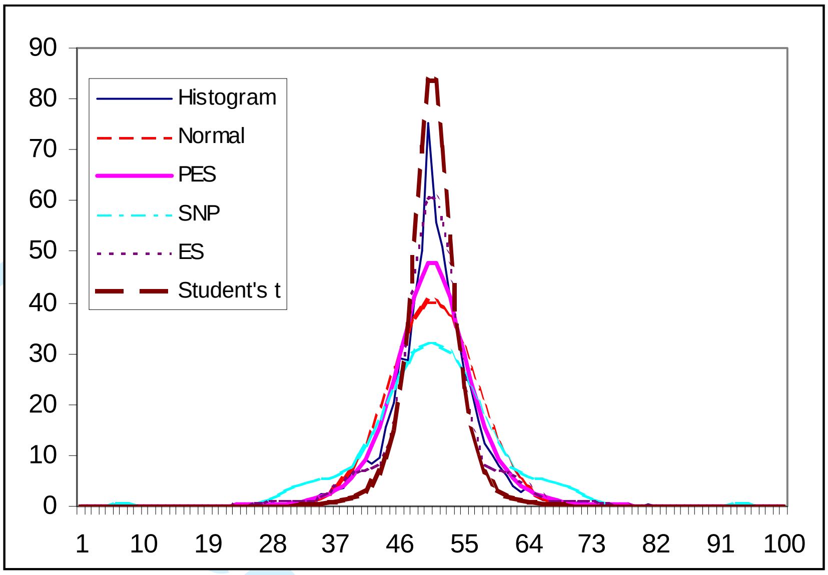


Figure 3.A. Fitted densities

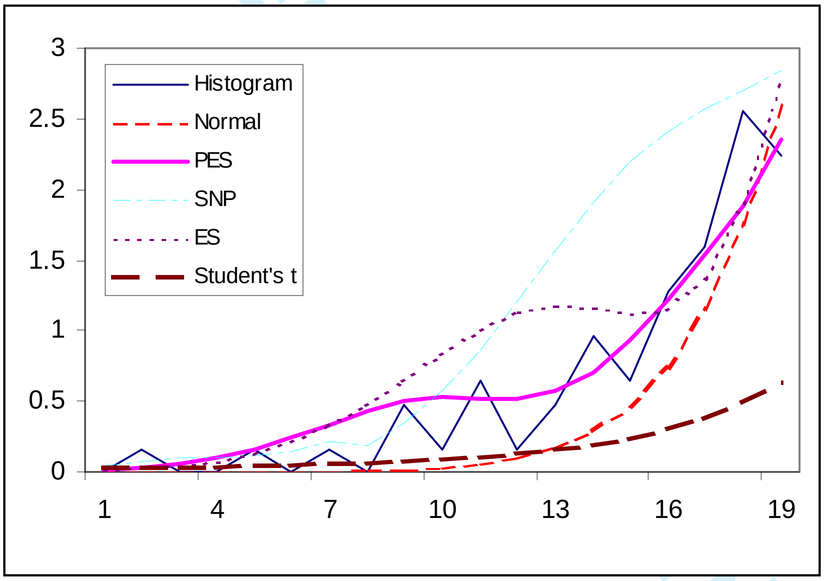


Figure 3.B. Fitted densities (left tails)

Figure 4. P-value plots of the PITs of $\widehat{u}_{it+1}$ obtained under the MGCII and MN models.

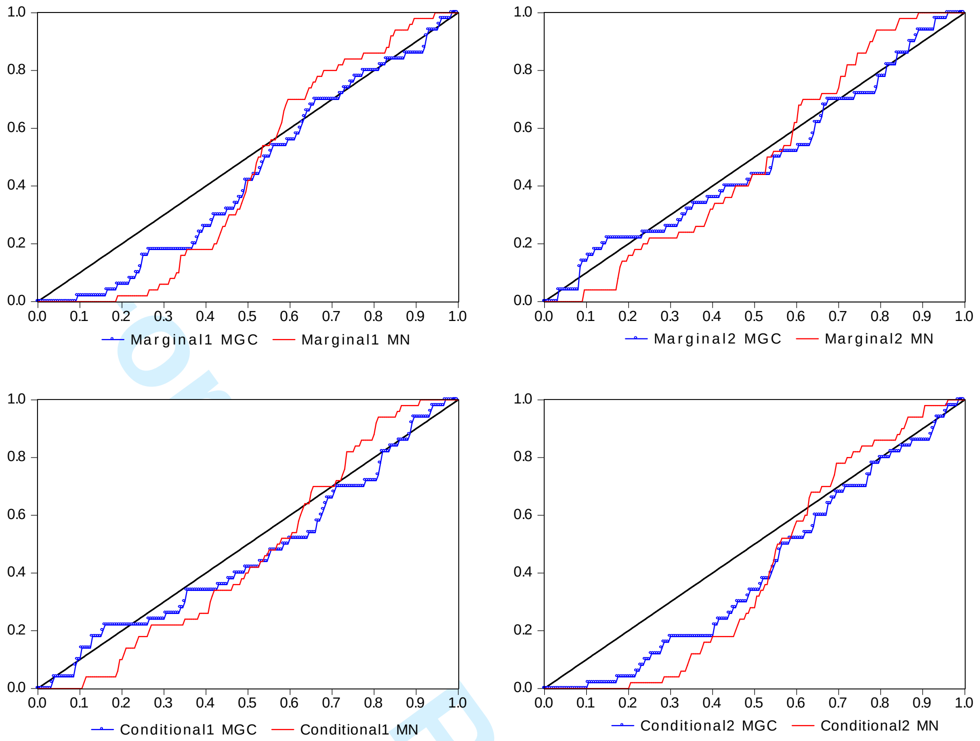
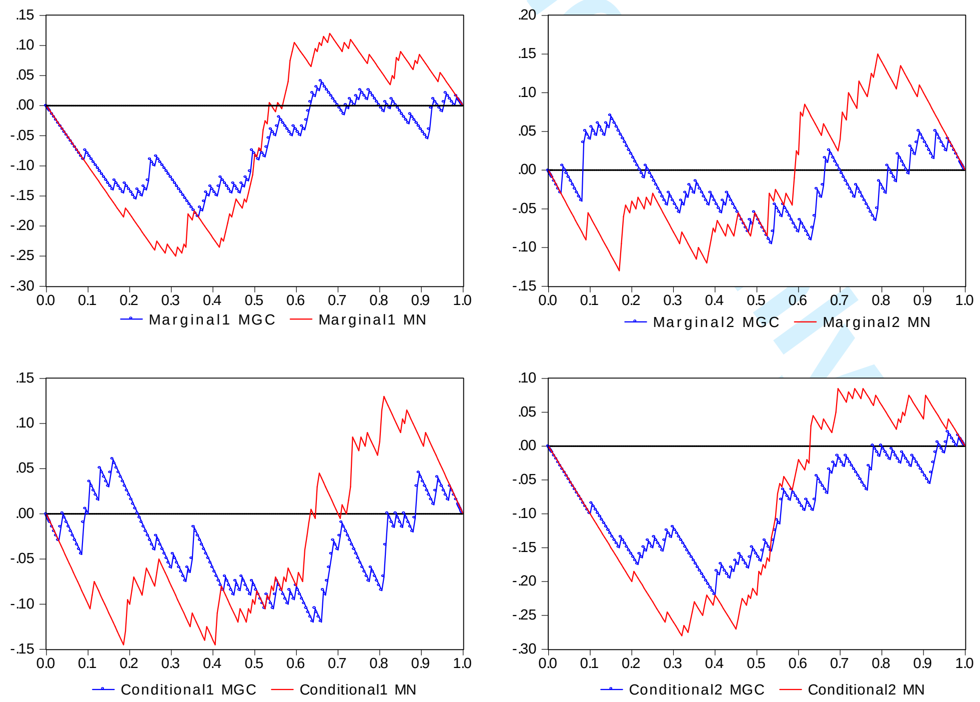


Figure 5: P-value discrepancy plots of the PITs of $\widehat{u}_{it+1}$ obtained under the MGCII and MN models.



References

[1] Aas, K., I. Haff, and X. K. Dimakos. 2006. Risk Estimation Using the Multivariate Normal Inverse Gaussian Distribution. *Journal of Risk*, 8, 39-59.

[2] Azzalini, A. and A. Dalla Valle. 1996. The Multivariate Skew Normal Distribution. *Biometrika*, 83, 715-726.

[3] Bauwens, L., S. Laurent. 2005. A New Class of Multivariate Skew Densities with Application to GARCH Models. *Journal of Business and Economic Statistics*, 23, 346-354.

[4] Bauwens, L., S. Laurent and J. V. K. Rombouts. 2005. Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics*, 21, 79-109.

[5] Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307-327.

[6] Bollerslev, T. 1990. Modelling the Coherence in Short-Run Nominal Exchange Rates: a Multivariate Generalized ARCH Model. *Review of Economics and Statistics*, 72, 498-505.

[7] Bollerslev, T. and J. Wooldridge. 1992. Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. *Econometric Reviews*, 11, 143-172.

[8] Boudt, K., C. Peterson and C. Roux. 2007. Estimation and Decomposition of Downside Risk for Portfolios with Non-normal Returns. Leuven University working paper No. 32.

[9] Corrado, C. J. and T. Su. 1996. Skewness and Kurtosis in S&P 500 Index Returns Implied by Option Prices. *Journal of Financial Research*, 19, 175-192.

[10] Corrado, C. J. and T. Su. 1997. Implied Volatility Skews and Stock Returns Skewness and Kurtosis Implied by Option Pricing. *European Journal of Finance*, 3, 73-85.

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

[12] Diebold, F. X., T. A. Gunther and S. A. Tay. 1998. Evaluating Density Forecasts with Applications to Financial Risk Management. *International Economic Review*, 39, 863-883.

[13] Diebold, F. X., J. Hahn and S. A. Tay. 1999. Multivariate Density Forecasts Evaluation and Calibration in Financial Risk Management: High-Frequency Returns of Foreign Exchange. *Review of Economics and Statistics*, 81, 661-673.

[14] Dittmar, R. F. 2002. Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns. *Journal of Finance*, 57, 369-403.

[15] Engle, R. F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflations. *Econometrica*, 50, 987-1007.

[16] Engle, R. F. 2002. Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. *Journal of Business and Economic Statistics*, 20, 339-350.

[17] Engle, R. F. and K. Sheppard. 2001. Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. *NBER* Working Paper No. 8554.

[18] Fenton, V. and A. R. Gallant.1996. Qualitative and Asymptotic Performance of SNP Density Estimators. *Journal of Econometrics*, 74, 77-118.

[19] Fiorentini, G., E. Sentana and G. Calzolari. 2003.Maximum Likelihood Estimation and Inference in Multivariate Conditional Heteroskedastic Dynamic Regression Models with Student t Innovations. *Journal of Business and Economic Statistics*, 21, 532-546.

[20] Gallant, A. R. and D. Nychka. 1987. Seminonparametric Maximum Likelihood Estimation. *Econometrica*, 55, 363-390.

[21] Gallant, A. R. and G. Tauchen. 1989. Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications. *Econometrica*, 57, 1091-1120.

[22] Harvey, C. R. and A. Siddique. 1999. Autoregressive Conditional Skewness. *Journal of Financial and Quantitative Analysis*, 34, 465-487.

[23] Harvey, C. R. and A. Siddique. 2000. Conditional Skewness in Asset Pricing Tests. *Journal of Finance*, 55, 1263-1295.

[24] Harvey, A. C., E. Ruiz and N. Sheppard. 1992. Unobservable Component Time Series Models with ARCH Disturbances. *Journal of Econometrics*, 52, 129-158.

[25] Jondeau, E. and M. Rockinger. 2001. Gram-Charlier Densities. *Journal of Economic Dynamics and Control*, 25, 1457-1483.

[26] Jondeau, E. and M. Rockinger. 2005. Conditional Asset Allocation under Non-Normality: How Costly Is the Mean-Variance Criterion? EFA 2005 Moscow Meetings Paper.

[27] Jondeau, E. and M. Rockinger. 2006. Optimal Allocation under Higher Moments. *European Financial Management*, 12, 29-55.

[28] Jondeau, E. and M. Rockinger. 2006. The Impact of News on Higher Moments. Swiss Finance Institute Research Paper No. 28.

[29] Jondeau, E. and M. Rockinger. 2007. The Economic Value of Distributional Timing. Swiss Finance Institute Research Paper No. 06-35.

[30] Jurczenko, E., B. Maillet and B. Negrea. 2004. A Note on Skewness and Kurtosis Adjusted Option Pricing Models under Martingale Restriction. *Quantitative Finance*, 4, 479-488.

[31] Kendall, M. and A. Stuart. 1977. *The Advanced Theory of Statistics*, Vol. I, 4th edn., Griffin & Co, London.

[32] Kotz, S. and S. Nadarajah. 2004. *Multivariate T-Distributions and Their Applications*, Cambridge University Press.

[33] León, A., G. Rubio and G. Serna. 2005. Autoregressive Conditional Volatility, Skewness and Kurtosis. *Quarterly Review of Economics and Finance*, 45, 599-618.

[34] León, A., J. Mencia and E. Sentana. 2007. Parametric properties of semi-nonparametric distributions, with applications to option valuation. Working Paper, Spanish Central Bank.

[35] Malevergne, Y. and D. Sornette. 2004. Multivariate Weibull Distributions for Asset Returns: I. *Finance Letters*, 2, 16-32.

[36] Mandelbrot, B. 1963. The Variation of Certain Speculative Prices. *Journal of Business*, 36, 1279-1313.

[37] 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.

[38] Morgan, J. P. 1996. RiskMetrics - Technical Document, Fourth Edition. New York: J. P. Morgan.

[39] Ñíguez, T. M. and J. Perote. 2004. Forecasting the Density of Asset Returns. STICERD Econometrics Discussion Paper 479. *London School of Economics*.

[40] Olcay, A. 2005. A New Class of Multivariate Distributions: Scale Mixture of Kotz-Type Distributions. *Statistics and Probability Letters*, 75, 18-28.

[41] Perote, J. 2004. The Multivariate Edgeworth-Sargan Density. *Spanish Economic Review*, 6, 77-96.

[42] Sahu, S. K., D. K. Dey and M. D. Branco. 2003. A New Class of Multivariate Skew Distributions with Applications to Bayesian Regression Models *Canadian Journal of Statistics*, 31, 129-150.

[43] Sargan, J. D. 1975. Gram-Charlier Approximations Applied to t Ratios of k-Class Estimators. *Econometrica*, 43, 327-347.

[44] Sargan, J. D. 1976. Econometric Estimators and the Edgeworth Approximation. *Econometrica*, 44, 421-448.

[45] Vlaar, P. J. G. and F. C. Palm. 1993. The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroskedasticity and Jumps. *Journal of Business and Economic Statistics*, 11, 351-360.