Gram–Charlier densities: a multivariate approach

ESTHER B. DEL BRIO†, TRINO-MANUEL ÑIGUEZ‡ and JAVIER PEROTE*§

†Department of Business and Finance, University of Salamanca, 37071 Salamanca, Spain
‡Department of Economics and Quantitative Methods, Westminster Business School, London NW1 5LS, UK
§Department of Economics, University Rey Juan Carlos, 28032 Madrid, Spain

(Received 16 March 2007; in final form 12 January 2009)

This paper introduces a new family of multivariate distributions based on Gram–Charlier and Edgeworth expansions. This family encompasses many of the univariate semi-non-parametric 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-non-parametric 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-of-sample framework for financial returns data. Our results show that the proposed specifications provide a reasonably good performance, and would therefore be of interest for applications involving the modelling and forecasting of heavy-tailed distributions.

Keywords: Empirical finance; Econometrics of financial markets; Financial assets; VaR

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) measures or short-fall probabilities, for which normality or non-normality is a key assumption. Since Mandelbrot (1963), financial econometricians have analysed the misspecification of models 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 in this line of research are provided by, for instance, Jurecenko et al. (2004), Jondeau and Rockinger (2005, 2006a, b, 2007) and Boudt et al. (2007). These authors provide evidence for 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, especially 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) have undergone important extensions since the constant conditional correlation (CCC) model of Bollerslev (1990) to the dynamic conditional correlation (DCC) model of Engle and Sheppard (2001) and Engle (2002); see Bauwens et al. (2005) for a complete survey of MGARCH models. On the other hand, different multivariate distributions have been introduced in financial econometrics, of 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, Bauwens and Laurent 2005), 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) 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-non-parametric (SNP) approach has the advantage of its general and flexible structure, since it endogenously admits as many 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, have not usually considered expansions beyond fourth order (see, e.g., Corrado and Su 1996, 1997, Harvey and Siddique 1999, Jondeau and Rockinger 2001 and León et al. 2005). It should be mentioned that the application of SNP densities requires 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 second and the third are also useful for out-of-sample analysis (see León et al. 2005 and Níguez and Perote 2004 for applications of the latter method in the in- and out-of-sample contexts). However, it is known that parametric constraints may lead to sub-optimization, with the model losing some of its flexibility, whilst reformulations may lead to theoretically less tractable specifications. Empirical work 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 (ES) distribution to the \(n\)-dimensional case, defining the Multivariate ES (MES), and provided evidence for 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 these 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) densities of León et al. (2005, 2009) and Níguez and Perote (2004), respectively. The theoretical properties of these densities (e.g. marginal distributions, cumulative distribution functions (cdf), 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 with the MES, the Multivariate Normal (MN) and the (skewed) Multivariate Student’s \(t\) (Sk-MST) through an empirical application to stock returns. We provide evidence 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 (in particular, 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 of 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 wide 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 outperforms 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 \(\rho_{ij},\ \forall i, j = 1, \ldots, n, \ i \neq j\)), and the so-called Hermite polynomials, \(H(\cdot)\), as given in the definition below.

Definition 2.1: A random vector \(X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\) belongs to the MGC family of distributions if it is distributed according to the following density function:

\[
F(X) = \frac{1}{n+1}G(X) + \frac{1}{n+1}\left\{\prod_{i=1}^{n}g(x_i)\right\}\left\{\sum_{i=1}^{n}\frac{1}{c_i}h(x_i)^{\text{A}_i}h(x_i)\right\},
\]

\(\text{(1)}\)

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

\(\ddagger\)Note that although we define the ‘standardized’ 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.
For this density, and based on the well-known orthogonality properties of Hermite polynomials, it can be straightforwardly proved that:

(i) the constant that weights the squared sum of Hermite polynomials for every variable $x_i$ (see the first proof in appendix A) is

$$c_i = \int g(x_i) \left[ 1 + \sum_{j=1}^{q} d_{ij} H_j(x_i) \right]^2 \, dx_i = 1 + \sum_{j=1}^{q} d_{ij} s_j,$$

(ii) the density integrates up to one (see the second proof in appendix A); and

(iii) the marginal density for variable $x_i$ is a mixture of a univariate Normal and a univariate SNP density of the type recently analysed by León et al. (2005), as shown by (see the third proof in appendix A)

$$f_t(x_i) = \frac{n}{n+1} g(x_i) + \frac{1}{n+1} \sum_{l=1}^{q} \left[ 1 + \sum_{j=1}^{q} d_{ij} H_j(x_i) \right] g(x_i).$$

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. 2009 for a complete description of the moments of the SNP density). This fact also permits us 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, the following equation considers conditional skewness for every variable $i$, $s_{it}$, in the ‘standardized’ MGCI expanded to the third term:

$$F_3(X) = \frac{1}{n+1} G(X) + \frac{1}{n+1} \left\{ \sum_{i=1}^{n} g(x_i) \right\} \times \left\{ \sum_{i=1}^{n} \left[ 1 + \sum_{j=1}^{q} d_{ij} H_j(x_i) \right]^2 \right\}.$$
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 $H_4(x_i) = x_i^4 - 6x_i^2 + 3$.

<table>
<thead>
<tr>
<th></th>
<th>ES</th>
<th>SNP</th>
<th>PES</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[x_i]$</td>
<td>0</td>
<td>$48d_3^2d_4$</td>
<td>0</td>
</tr>
<tr>
<td>$E[x_i^2]$</td>
<td>1</td>
<td>$1 + 42d_3^2 + 216d_4^2$</td>
<td>$1 + 42d_3^2 + 216d_4^2$</td>
</tr>
<tr>
<td>$E[x_i^3]$</td>
<td>6$d_3$</td>
<td>$12d_3^2 + 576d_3d_4$</td>
<td>0</td>
</tr>
<tr>
<td>$E[x_i^4]$</td>
<td>$3 + 24d_3$</td>
<td>$3 + 450d_3^2 + 2952d_4^2 + 48d_4$</td>
<td>$3 + 450d_3^2 + 2952d_4^2 + 48d_4$</td>
</tr>
</tbody>
</table>

Normal and the GC, but also on the correlation between them. 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 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 provide 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, in addition to being capable of capturing multimodality, which makes them very useful for accurately forecasting risk measures related to the tails of assets returns distributions.

Furthermore, the MGCI distribution overcomes the aforementioned non-positivity problem that may arise when estimating the MES density of Perote (2004).†

$$F_{ES}(X) = G(X) + \left\{ \prod_{i=1}^{n} g(x_i) \right\} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{q} d_{ik} H_i(x_i) \right\}. \quad (14)$$

The second case that is noteworthy arises when $A_i = \text{diag}(1, d_{i1}, \ldots, d_{i8})$, $\forall i, j = 1, \ldots, n$. For these Hermite polynomial weighting matrices, the resulting density is defined in equation (15), which we denote as MGCI:

$$F_{II}(X) = \frac{1}{n+1} G(X) + \frac{1}{n+1} \left\{ \prod_{i=1}^{n} g(x_i) \right\} \times \left\{ \sum_{i=1}^{n} \frac{1}{c_i} \left[ 1 + \sum_{j=1}^{q} d_{ijk} H_i(x_i)^2 \right] \right\}. \quad (15)$$

Obviously, this density is a particular case of the former formulation but it may result in being 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 (16), are mixtures of a univariate Normal and the univariate PES defined by Níguez and Perote (2004)

$$f_{II}(x_i) = \frac{n}{n+1} g(x_i) + \frac{1}{(n+1)c_i} \left[ 1 + \sum_{j=1}^{q} d_{ijk} H_i(x_i)^2 \right] g(x_i). \quad (16)$$

Therefore, the MGCI distribution moments can be obtained as a combination of those of the univariate Gaussian and PES. In particular, the $k$th order PES even moment can be expressed as

$$m_{ik} = \frac{1}{c_{i}} E_N[x_i^k] + \frac{1}{c_{i}} \sum_{j=1}^{n} \sum_{j=0}^{k/2} \beta_{ijk} d_{ij}^2, \quad \forall k \text{ even.} \quad (17)$$

where $E_N[x_i^k]$ denotes the $k$th order moment of the Gaussian density and $\{\beta_{ijk}\}_{k=0}^{\infty}$ is the sequence of constants that makes $x_i^k = \sum_{j=0}^{\infty} \beta_{ijk} H_i(x_i)^2$ (see Níguez and Perote (2004) for the details of the moments of this distribution). For the sake of clarity, table 1 includes the first four moments of the ‘standardized’ PES expanded to fourth order compared with the ES and SNP counterparts.

Regarding the cross-moments, all the comments stated for MGCI also apply for MGCI. Furthermore, the MGCI cdf can easily be worked out as shown in equation (18) (see the fifth proof in appendix A), and, consequently, they can be used for risk management purposes, either for modelling and forecasting credit risk, portfolio VaR or short-fall probabilities. The multivariate cdf of MGCI can be obtained analogously in terms of the cdf of the univariate $N(0, 1)$ and

†Note that, for the maximum likelihood estimates, the MES must necessarily be 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).
univariate SNP distributions (see León et al. (2009) for further details)
\[
\Pr[x_i \leq a_1, \ldots, x_n \leq a_n] = \frac{1}{n!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(x_1) \cdots d x_n + \sum_{j=1}^{n-1} \int_{-\infty}^{a_i} g(x_j) d x_j 
\times \prod_{i=1}^{n} \left[ \int_{-\infty}^{a_i} g(x_i) d x_i - \frac{g(a_i)}{c_i} \sum_{j=1}^{q} \sum_{k=0}^{s-1} \frac{s!}{(s-k)!} \right] 
\times H_{s-k}(a_i) H_{s-k-1}(a_i)
\]

The MGC densities straightforwardly admit the specification of GARCH-type processes (Engle 1982, Bollerslev 1986) to explain the dynamics of their conditional moments. In particular, the conditional variances, \(k_n^2\), are introduced by considering transformations of the type \(u_t = A(x) \cdot X_t\), where \(A(x) = diag(k_1, k_2, \ldots, k_n)\). Specifically, we consider the following specification:
\[
r_t = \mu_t(\Phi) + u_t,
\]

In the next section we test the performance of the bivariate versions of the MGCI and MGCH 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, \(\nu\). In addition, we also include a comparison of the asymmetric versions of the MES and MGCI densities (Sk-MES and Sk-MGCI, respectively) with the skewed MST (Sk-MST) of Bauwens and Laurent (2005). The ‘standardized’ cases of the \(n\)-dimensional MST and Sk-MST are defined in equations (21) and (22), respectively:
\[
F_{ST}(X | \nu) = \frac{\Gamma((v+n)/2)}{(\pi(v-2))^{n/2} \Gamma(v/2)} \left[ 1 + \frac{XX}{v-2} \right]^{-(v+n)/2},
\]
\[
F_{SST}(X | \nu, \lambda) = \frac{\Gamma((v+n)/2)}{(\pi(v-2))^{n/2} \Gamma(v/2)} \left[ \prod_{i=1}^{n} \frac{2b_i}{\lambda_i + (1/\lambda_i)} \right] 
\times \left[ 1 + \frac{X'X}{v-2} \right]^{-(v+n)/2},
\]

where \(\Gamma(\cdot)\) is the gamma function and \(\lambda = (\lambda_1, \ldots, \lambda_n) > 0\) is the vector of asymmetry parameters; therefore, \(F_{SST}(X | \nu, \lambda)\) is skewed to the right (left) if \(\ln(\lambda_i) > 0\) (<0) and \(F_{SST}(X | \nu, \lambda)\) reduces to \(F_{ST}(X | \nu)\) when \(\lambda = 1\), and \(\nu\) is constrained to be larger than 2 to ensure the existence of the covariance matrix.

3. An empirical application to portfolio returns

The data used are daily returns of the S&P 500 and the Hang-Seng indexes of the New York and Hong Kong Stock Exchange, respectively, \(r_t = (r_{t1}, r_{t2})\), over the period December 19, 1991 to December 18, 2006 for a total of \(T = 3913\) observations, obtained from Datastream. Plots and descriptive statistics of \(r_t\) are presented in figure 1.

Let the conditional distribution of \(r_t\) be either MN, MES, MGCI, MGCHI or MST, with the conditional mean and covariance matrix modelled according to equations (19) and (20). 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 \(r_t\), and a GARCH(1, 1) process to account for volatility clustering in the conditional variance of \(r_t\):
\[
\mu_{it} = \phi_{0} + \phi_{1} r_{i,t-1}, \quad \forall i = 1, 2,
\]

The estimation procedure is carried out in two steps using an in-sample window of \(S = 3512\) observations. Firstly, the AR(1) process is estimated by ordinary least squares and, secondly, covariance matrix coefficients and density function parameters are estimated by (quasi-)maximum likelihood (QML) using the AR(1) residuals from the first step. Robust QML covariance estimators are calculated using the formula of Bollerslev and Wooldridge (1992). The Hermite polynomial expansions were truncated at the eighth term according to accuracy criteria. For the purpose of concentrating on the heavy...
tails of the distribution and considering the fact that the odd parameters were not found to be statistically jointly significant, in a first approach these odd parameters were constrained to zero.†

The likelihood function is maximized using the Newton–Raphson method. We observe that estimation of the MGC models is not computationally very demanding provided that starting values are chosen properly. As MGC models are nested, a usual procedure for choosing 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 function 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 does 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 optima.‡

Table 2 displays the estimates and the 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 ($\phi_{10} > \phi_{11}$), and the AR(1) slope coefficient is significantly higher for the S&P 500. All AR(1) coefficients are approximately non-significant at the 5% level, but $\phi_{10}$, $\phi_{20}$, and $\phi_{11}$ are at the 10% level. Secondly, we observe the typical estimates of GARCH processes for financial returns; for both indexes the GARCH parameter estimates of all models reflect the existence of clustering and a high persistence of volatility ($\alpha_1 + \alpha_2$ is close to, but smaller than, one), although the sum ($\alpha_1 + \alpha_2$) is significantly lower for the MCGII model, in line with the results of Niguez and Perote (2004).

In relation to the estimated correlations, $\hat{\rho}$, they are also significant and slightly higher for the MGC and MES models. It should be noted though that the correlation coefficients and the conditional variance parameters of those specifications have to be interpreted carefully. For example, $\rho$ 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, for the analysed MGC densities, the Pearson’s correlation

†The corresponding likelihood ratio (LR) test accepted the null hypothesis $H_0: d_0 = d_1 = d_2 = 0$ for all the considered densities. The LR test results are not displayed for the sake of simplicity, but are available from the authors upon request.

‡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 becoming trapped in successive local optima.
Bivariate density for S&P 500 (variable 1) and Hang-Seng (variable 2) indexes. $z$-ratios in parentheses. The asterisk denotes approximate non-significance at the 5% confidence level.

Further, the stationarity conditions of the GARCH processes in MGC distributions are also slightly different from the usual ones.\(^\dagger\)

In relation to the parameter weight estimates in table 2 ($d_{ij}$, $i = 1$ (S&P 500), $i = 2$ (Hang-Seng), $s = 2, 4, 6, 8$) (hereafter $i,j = 1,2$) we observe that most of them are significant at reasonable confidence levels, with only $d_{12}$ clearly not significant for the MGCI model and $d_{12}$ for all SNP models. These estimates show that the portfolio distribution is highly leptokurtic and that the MGC models are able to capture that tail shape parsimoniously. This result is confirmed by the estimated degrees of freedom of the MST model, $\nu$, which equals 7.1. An interesting observation is that the coefficient $d_{26}$ is clearly significant in most cases, reinforcing the fact that the densities need to be expanded at least up to the eighth 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 larger the $j$th subindex.

For the purpose of comparing the accuracy of the different specifications and as a first orientation approach, table 2 includes the log-likelihood value ($\ln L$) and the BIC, computed as $-\ln L + p\ln n(S)/2$, where $p$ is the number of parameters of the model. According to these criteria the densities based on Edgeworth and GC expansions out-perform the most popular distributions in finance (MN and MST). This improvement in accuracy is due to the fact that MGC models present greater flexibility than MST models since they operate with more parameters to parsimoniously account for the shape of the target distribution. 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 MGCI density in comparison with the MN for their different ranges and domains, including the multimodality feature, is provided in figure 2. Figure 2(a), (c) and (e) correspond to the fitted MGCI and figure 2(b), (d) and (f) to the MN. In particular, figures 2(a) and (b) present the whole domain of the functions, and the remaining figures illustrate details of the distribution tails. It is noteworthy that the MGCI 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 MGCI captures more accurately the leptokurtic density behaviour since it assigns positive probabilistic mass to areas in the tails where the MN does not (see figures 2(e) and (f)).

\(^\dagger\)See Níguez 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&P 500 under different specifications (MN, MST, MES, MGCI and MGCII) in comparison with the histogram of the data. Figure 3(a) presents the densities for the whole domain, whilst figure 3(b) includes only the left tails of the distributions. From these figures 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-non-parametric alternatives (MES or MGCI).† Therefore, in the out-of-sample application below we analyse the performance

†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 might explain misspecification for tail behaviour of MST models.
of the MGCII as a representative well-behaved MGC distribution.

The aforementioned comparisons are focused on the tail behaviour of symmetric density specifications. Nevertheless, financial returns also seem to feature skewness. This was especially found when conditional skewness processes were incorporated into the modelling of the returns distribution (Harvey and Siddique 1999, 2000). In order to also account for this feature we estimated the Sk-MES and Sk-MGCI distributions, expanded up to the eighth 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.

<table>
<thead>
<tr>
<th>Table 3. Multivariate skewed densities.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sk-MST</td>
</tr>
<tr>
<td>( \alpha_{10} )</td>
</tr>
<tr>
<td>( \alpha_{11} )</td>
</tr>
<tr>
<td>( \alpha_{20} )</td>
</tr>
<tr>
<td>( \alpha_{21} )</td>
</tr>
<tr>
<td>( \gamma_{10} )</td>
</tr>
<tr>
<td>( \gamma_{20} )</td>
</tr>
<tr>
<td>( \gamma_{21} )</td>
</tr>
<tr>
<td>( \gamma_{22} )</td>
</tr>
<tr>
<td>( \mu_{12} )</td>
</tr>
<tr>
<td>( \mu_{23} )</td>
</tr>
<tr>
<td>( \mu_{24} )</td>
</tr>
<tr>
<td>( \mu_{25} )</td>
</tr>
<tr>
<td>( \mu_{26} )</td>
</tr>
<tr>
<td>( \mu_{27} )</td>
</tr>
<tr>
<td>( \mu_{28} )</td>
</tr>
<tr>
<td>( \rho )</td>
</tr>
<tr>
<td>( \psi )</td>
</tr>
<tr>
<td>Ln L</td>
</tr>
<tr>
<td>BIC</td>
</tr>
</tbody>
</table>

Bivariate density for S&P 500 (variable 1) and Hang-Seng (variable 2) indexes. \( \gamma \)-Ratios in parentheses. The asterisk denotes approximate non-significance at the 5% confidence level.

The AR(1) coefficients of the models in table 3 are not presented for simplicity, since they are the same as those in table 2 because of the two-step estimation procedure.
provide more evidence for 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:

\[ s_t = \gamma_0 + \gamma_1 \left( \frac{u_{it-1}}{y_{it-1}} \right)^3 + \gamma_2 s_{it-1}, \]

where \(-1 < \gamma_0 < 1\) represents the unconditional skewness, and \(\gamma_1, \gamma_2 \in \mathbb{R}\). The second coefficients and degrees of freedom, are similar to those of the conditional variance processes, and the correlation coefficients in table 2. A second density, which we denote CSk-MGCI, are also displayed where

\[ \frac{1}{C13} \]

and

\[ \frac{1}{C21} \]

are given in table 3. Firstly, we note that the estimates of the conditional variance processes, and the correlation coefficients of the corresponding symmetric models in table 2. A second observation is that, for the S&P 500 index, the skewness coefficient, \(\gamma_1\), in the Sk-MST model is not significantly different from 1, indicating that the S&P 500 returns density is unconditionally symmetric, whilst for the Hang-Seng, \(\gamma_2\) is significantly smaller than 1 at the 10% level, indicating that the Hang-Seng returns density is slightly unconditionally skewed. This result is confirmed by the individual significance tests of the coefficients \(\gamma_1\) and \(\gamma_2\) in the Sk-MGCI model, and the even weights coefficients in the Sk-MES and Sk-MGCI models. However, 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 the Sk-MES and Sk-MGCI distributions. An explanation for 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 \(\gamma_1\) and \(\gamma_2\) are positive and significant, which indicates 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 their persistence. These 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 less than that observed in volatility, as the sum \(\gamma_1 + \gamma_2\) is not as close to one as the corresponding value 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 only illustrative since they are only comparable for the Sk-MES and Sk-MGCI models.

Finally, we test the performance of the MGC densities in forecasting the full density of the portfolio and compare the forecasts with those of a MN model using the methodology of 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, \(u_{it+1} = \left( r_{it+1} - \bar{r}_{it+1} \right) / \hat{k}_{it+1}\), through the out-of-sample period (\(N = 400\) observations). The resulting so-called probability integral transforms (PITs) sequences, labelled \(\hat{p}_{it}, \hat{p}_{ijt}\), \(\forall i, j = 1, 2\) are i.i.d. \(U(0,1)\) under the correct density specification

\[ p_t = \int_{-\infty}^{u_{it+1}} f_{i+j}(x)dx. \]

\[ p_{ijt} = \int_{-\infty}^{u_{ijt+1}} f_{i+j}(x)dx. \]

\[ = \int_{-\infty}^{u_{ijt+1}} \int_{-\infty}^{u_{ijt}} f_{i+j}(x)dx_{i+j+1}dx_{ijt+1} \]

\[ = \int_{-\infty}^{u_{ijt+1}} \int_{-\infty}^{u_{ijt}} f_{i+j}(x)dx_{i+j+1}dx_{ijt+1} \]

where \(f_1()\), \(f_{ij}()\) and \(f_2()\) denote marginal, conditional and joint distributions, respectively. Moreover, since \(p_t\) is also interpreted as the \(p\)-value corresponding to the quantile \(\hat{u}_{it+1}\) of the forecasted density we use the \(p\)-value plot methods of Davidson and MacKinnon (1998) to compare the forecasting performance of the models.† Therefore, if the model is correctly specified the difference between the cdf of \(p_t\) and the 45° line should tend to zero asymptotically. The empirical distribution function of \(p_t\) can easily be computed as

\[ \hat{p}_{it}(y) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(p_t \leq y). \]

where \(\mathbb{I}(p_t \leq y)\) is an indicator function that takes the value 1 if its argument is true and 0 otherwise, and \(y\) is an arbitrary grid of \(\hat{p}\) points, which is made finer on its extremes to highlight differences in the goodness-of-fit of the density tails of the models. Alternatively, the \(p\)-value discrepancy plot (i.e. plotting \(\hat{P}_{it}(y) - y\) versus \(y\)) can be more revealing when it is necessary to discriminate among specifications that perform similarly in terms of the \(p\)-value plot (Fiorentini et al. 2003). Consequently, under correct density specification, the variable \(\hat{P}_{it}(y) - y\) 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 MGCI (blue line) density. An immediate observation is that the Hang-Seng returns density is slightly different from 1, indicating that the S&P 500 returns density is unconditionally skewed. This result is confirmed by the individual significance tests of the coefficients \(\gamma_1\) and \(\gamma_2\) in the Sk-MGCI model, and the even weights coefficients in the Sk-MES and Sk-MGCI models. However, 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 the Sk-MES and Sk-MGCI distributions. An explanation for 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 \(\gamma_1\) and \(\gamma_2\) are positive and significant, which indicates 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 their persistence. These 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 less than that observed in volatility, as the sum \(\gamma_1 + \gamma_2\) is not as close to one as the corresponding value 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 only illustrative since they are only comparable for the Sk-MES and Sk-MGCI models.

Finally, we test the performance of the MGC densities in forecasting the full density of the portfolio and

\[ \hat{p}_{it}(y) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(p_t \leq y). \]

where \(\mathbb{I}(p_t \leq y)\) is an indicator function that takes the value 1 if its argument is true and 0 otherwise, and \(y\) is an arbitrary grid of \(\hat{p}\) points, which is made finer on its extremes to highlight differences in the goodness-of-fit of the density tails of the models. Alternatively, the \(p\)-value discrepancy plot (i.e. plotting \(\hat{P}_{it}(y) - y\) versus \(y\)) can be more revealing when it is necessary to discriminate among specifications that perform similarly in terms of the \(p\)-value plot (Fiorentini et al. 2003). Consequently, under correct density specification, the variable \(\hat{P}_{it}(y) - y\) 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 MGCI (blue line) density. An immediate observation that emerges from these plots is that the MGCI model provides reasonably good performance for forecasting the full density of the portfolio and clearly out-performs the MN model commonly used in financial risk applications.

4. Concluding remarks

This paper introduces a family of multivariate distributions based on Edgeworth and GC expansions.

†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 between alternative models according to their performance for forecasting the full density.
Figure 4. *p*-Value plots of the PITs of $\hat{u}_{t+1}$ obtained under the MGCI and MN models.

Figure 5. *p*-Value discrepancy plots of the PITs of $\hat{u}_{t+1}$ obtained under the MGCI and MN models.
This family encompasses most of the univariate densities proposed in the 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 precursors 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 peaks, 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 generalize 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 the model of Harvey and Siddique (1999).

Nevertheless, the good performance in terms of accuracy measures in the whole domain does not necessarily imply the best fit in the distribution tails. We show that, in some cases, other, more parsimonious specifications, such as MGCI, 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 up a hopefully fruitful line of research providing general formulations for MGC densities, and provides evidence of their reasonably good in- and out-of-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 the 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.

Acknowledgements

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

References


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

$$c_i = \int g(x_i)dx_i + \int \left[ \sum_{j=1}^{q} d_{ij} H_i(x_j) \right] g(x_i)dx_i$$

$$\quad + 2 \sum_{j=1}^{q} d_{ij} \int H_i(x_j) g(x_i)dx_i$$

$$\quad - \int \left[ \sum_{j=1}^{q} d_{ij} H_i(x_j) \right] g(x_i)dx_i$$

$$\quad = 1 + \sum_{j=1}^{q} d_{ij}^2 \int H_i(x_j) g(x_i)dx_i$$

$$\quad = 1 + \sum_{j=1}^{q} d_{ij}^2 \int H_i(x_j) g(x_i)dx_i = 1 + \sum_{j=1}^{q} d_{ij}^2!$$

**Proof:** The MGCI density integrates up to one provided that $c_i$ are the constants in the above proof:

$$\int \cdots \int f_i(x_1) \cdots dx_n = 1 + \sum_{i=1}^{n} c_i$$

**Appendix A: Proofs**

This appendix includes the proofs of some properties of the MGC densities. In particular, 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 MGCI are derived. The corresponding proofs for other multivariate densities of the same family can be obtained likewise.
\[
\begin{align*}
&= \frac{1}{n+1} g(x_i) + \frac{1}{n+1 c_i} g(x_i) \left[ 1 + \sum_{j=1}^{q} d_i H_i(x_i) \right]^2 \\
&+ \frac{n-1}{n+1} g(x_i) \\
&= \frac{n}{n+1} g(x_i) + \frac{1}{n+1 c_i} g(x_i) \left[ 1 + \sum_{j=1}^{q} d_i H_i(x_i) \right]^2.
\end{align*}
\]

**Proof:** The co-movements of the MGCI density can be obtained in terms of the corresponding co-movements of the MN density and the univariate moments of both the normal and SNP distribution. An example for an element of the co-kurtosis matrix, called the co-volatility between \(x_1\) and \(x_2\), is given as

\[
k_{1122} = \int \cdots \int x_1^2 x_2^2 F_1(x_1)dx_1 \cdots dx_n
\]

\[
= \frac{1}{n+1} \int \cdots \int x_1^2 x_2^2 G(x_1)dx_1 \cdots dx_n \\
\quad + \frac{1}{(n+1)} \int x_1^2 \frac{1}{c_i} g(x_i) \left[ 1 + \sum_{j=2}^{q} d_i H_i(x_j) \right]^2 dx_i \\
\times \int x_2^2 g(x_2)dx_2 \prod_{j=1}^{n} \left[ \int g(x_j)dx_j \right] \\
\quad + \frac{1}{(n+1)} \int x_1^2 g(x_1)dx_1 \int x_2^2 \frac{1}{c_2} g(x_2) \\
\times \left[ 1 + \sum_{j=2}^{q} d_2 H_2(x_j) \right]^2 dx_1 \prod_{j=1}^{n} \left[ \int g(x_j)dx_j \right] \\
\quad + \frac{1}{(n+1)} \sum_{j=1}^{n} \left[ \int x_1^2 g(x_1)dx_1 \int x_2^2 g(x_2)dx_2 \frac{1}{c_i} g(x_i) \\
\times \left[ 1 + \sum_{j=2}^{q} d_i H_i(x_j) \right]^2 dx_2 \prod_{j=1,j\neq 1,2}^{n} \left[ \int g(x_j)dx_j \right] \right].
\]

\[
= \frac{1}{n+1} \left[ E_{MN}[x_1^2 x_2^2] + E_{GC}[x_1^2] E_{SN}[x_2^2] \\
+ E_{SN}[x_1^2] E_{GC}[x_2^2] + (n-2) E_{SN}[x_1^2] E_{SN}[x_2^2] \right].
\]

**Proof:** The cdf of the MGCHI density can be obtained in terms of the cdf of the MN density and the cdf of the univariate normal and PES distribution:

\[
\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} F_{1D}(X)dx_1 \cdots dx_n \\
= \frac{1}{n+1} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} G(x)dx_1 \cdots dx_n \\
\quad + \frac{1}{n+1} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \left[ \prod_{j=1}^{n} g(x_j) \right] \\
\quad \times \left[ \sum_{i=1}^{q} \frac{1}{c_i} \left[ 1 + \sum_{j=1}^{q} d_i^2 H_i(x_i) \right]^2 \right] dx_i \cdots dx_n \\
= \frac{1}{n+1} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} G(X)dx_1 \cdots dx_n \\
\quad + \frac{1}{n+1} \sum_{j=1}^{n} \left[ \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} g(x_j) dx_j \left[ 1 + \sum_{i=1}^{q} d_i^2 H_i(x_j) \right]^2 \right] dx_i \\
\quad \times \prod_{j=1,j\neq i}^{n} g(x_j)dx_j \\
\quad + \frac{1}{n+1} \sum_{j=1}^{n} \left[ \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} g(x_j)dx_j - \frac{(a_i)}{c_i} \sum_{i=1}^{q} d_i^{2} \sum_{k=0}^{s-1} \frac{s!}{s-k} \\
\times H_{s-k}(a_i) H_{s-k-1}(a_j) \prod_{j=1,j\neq i}^{n} \int_{-\infty}^{x_n} g(x_j)dx_j \right].
\]

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