Multivariate Distributions based on General Moments Expansions: Evidence from Exchange Rates

Trino-Manuel Ñíguez

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

Javier Perote*

Department of Economics, University of Salamanca, Salamanca, Spain

Antonio Rubia

Department of Finance and Accounting, University of Alicante, Alicante, Spain

[Work in progress]

Abstract

This paper proposes multivariate Semi-Nonparametric distributions (SNP) based on the General Moments Expansion (GME) to model portfolio returns distribution. The multivariate GME is as flexible as other multivariate SNP distributions based on Gram-Charlier series and thus is capable of capturing salient empirical regularities of financial data but it present at least two important advantages: (i) it embodies a much simpler polynomial structure which simplifies the analytical tractability of the density, specially when positive transformations are implemented; and (ii) it straightforwardly admits the consideration of any non-Normal distribution used as basis with the only requirement of having finite moments up to the expansion order. We show that if the expansion uses the Gaussian distribution as basis the two-step estimation procedure introduced by Engle (2002) can also be formally implemented for the GME distribution, thus overcoming the "curse of dimensionality" of multivariate volatility modeling. We compare the performance of different multivariate SNP alternatives showing that the GME-DCC model is very easy to implement and provides very accurate results for capturing portfolio returns distribution.

JEL classification: C16, G1.

Keywords: Semi-nonparametric methods; multivariate Non-normal distributions, DCC model, GARCH models, Financial markets.

*Correspondence to: Javier Perote, Departamento de Economía e Historia Económica, Facultad de Economía y Empresa (Edif. FES), Campus de Miguel de Unamuno, Universidad de Salamanca, 37007 Salamanca, Spain. Tel: +34 923 294400 (ext. 3512). E-mail: perote@usal.es

1 Introduction

The modeling of the multivariate volatility dependences in a mean-variance portfolio allocation framework has been traditionally tackled by different parametric specifications introduced to simplify the computation of the covariance matrix of returns and at the same time to guarantee positive definiteness. For these purposes there have been proposed multivariate extensions of GARCH models (e.g. the diagonal vech model (Bollerslev, Engle & Wooldridge, 1988) or the BEKK (Engle & Kroner, 1995)), Factor models (Engle, Ng & Rothschild, 1990, and Sentana & Fiorentini, 2001) that decompose common and idiosyncratic factors in portfolio dynamics or two-step estimation models (e.g. the Orthogonal GARCH (Alexander & Chibumba, 1996, and Van der Weide, 2002), the Constant Conditional Correlation (CCC) (Bollerslev, 1990), the Dynamic Conditional Correlation (DCC) (Engle, 2002, and Engle & Sheppard , 2001) or the Dynamic Equicorrelation (DECO) model (Engle & Kelly, 2007)) (see Bauwens, Laurent & Rombouts (2005) for a comprehensive review on multivariate volatility models).

On the other hand the evidence of leptokurtosis and skewness found in portfolio returns suggests the consideration of other multivariate distributions than the Normal. In order to incorporate such features copulas (Chen, Fan & Tsyrennikov, 2006) and different parametric multivariate distributions (skewed Normal (Azzalini & Dalla Valle, 1996), Student's t (Kotz & Nadarajah, 2004), Weibull (Malevergne & Sornette, 2004) or Kotz-type (Olcay, 2005)) may be used. The copula's approach provides a straightforward generalization of any univariate marginal distributions to a multivariate framework, but those models become analytically intractable for different financial applications that require highly computationally demanding numerical algorithms. Furthermore the existing heavy-tailed and skewed multivariate distributions may not be flexible enough to capture some of the salient features of financial returns. Alternatively the Semi-Nonparametric (SNP) approach seems to be a flexible an easy to implement method that can solve all these shortcomings. Particularly, Del Brio, Ñíguez & Perote (2009) have shown that multivariate densities based on Gram-Charlier (GC hereafter) series accurately account for not only fat tails and skewness but also different jumps on the probabilistic mass at the tails of density returns by means of a flexible and parsimonious representation. This potential advantage of the SNP distributions lies in the fact that, and under regularity conditions, any frequency function may be represented in terms of (infinite) GC series. Moreover, Del Brio, Ñíguez & Perote (2010) provide a generalization of the two-step DCC method to the SNP densities, which up to our knowledge seem to be the first non-Normal distributions that formally admit this two-step procedure. The two-step method overcomes the so-called "curse of dimensionality" by the Quasi-Maximum Likelihood estimation (QMLE) of the conditional variances parameters under univariate Normal distributions for every dimension and the rest of the density parameters by Limited Information Maximum Likelihood (LIML) estimation under the SNP specification concentrated on the estimates of the first step. For these reasons the multivariate SNP distributions seem to be a useful tool to fit and forecast portfolio risk.

Nevertheless the SNP approach also involves several shortcomings mainly due to the fact that the GC expansions must be truncated in practice and thus positiveness needs to be imposed by transformations of the Gallant & Nychka (1987) and Gallant & Tauchen (1989) type. This solution is far from being innocuous since it affects the statistical properties of the distribution and, in particular, introduces non-linear restrictions among all the distribution moments. This fact may complicate the implementation of these multivariate distributions and make them less appealing from the practitioners viewpoint. In this paper we introduce a new multivariate SNP family of distributions, which we name Multivariate General Moments Expansion (MGME hereafter), that preserves the good properties of the multivariate GC (MGC hereafter) distributions but presents a much simpler formulation and therefore are more straightforward to implement for empirical purposes. The new multivariate family of distributions generalizes the GME proposed in Ñíguez & Perote (2007) to the multivariate framework and consequently inherits all the good properties of the GME: *(i)* generality: the MGME only requires that the distribution used as basis has finite moments up to the length of the expansions and MGC distributions can be obtained by reformulations in terms of the Hermite polynomials, *(ii)* positiveness: sufficient conditions or transformations to ensure it can be straightforwardly proposed, *(iii)* empirical tractability: it presents an extremely simple formulation and if the Gaussian distribution is used as basis then it admits the decomposition proposed by Engle (2002), which eases the model implementation, and *(iv)* it yields a reasonable empirical performance for capturing the density of portfolio returns.

The remainder of the article is divided into the following three sections. In Section 2, we define the MGME distribution and discuss its theoretical properties, including the two-step estimation of the Normal-GME distributed portfolio model. Section 3 provides an empirical application to two bivariate portfolios composed of exchange rates (FX hereafter) returns, to test the performance of the proposed model. Section 4 summarizes the main conclusions.

2 The multivariate GME distribution

The modeling of the multivariate distribution of financial variables faces serious obstacles that are still not fully resolved. For instance, the generalization of univariate marginal distributions to a multivariate framework has been successfully achieved by the use of copulas, but those models present drawbacks related with the integration of the joint distribution (e.g., to compute moments), which becomes analytically intractable and requires the implementation of highly computationally demanding numerical algorithms (see Jondeau, Poon & Rockinger, 2007, p. 196). On the other hand, the use of the existing heavy-tailed and skewed multivariate distributions are interesting alternatives, but in most cases those distributions cannot capture some of the salient features of financial returns due to the lack of a sufficiently flexible parametric structure. Moreover, the multivariate time-varying models can also be implemented for capturing conditional moments, but at the cost of a parameter structure that might give place to very complicated specifications for large-dimensional portfolios. That problem has been tackled with more parsimonious multivariate models and through different estimation procedures such as the method of moments (Polanski & Stoja, 2010) or estimation in two stages as in the DCC model. Nevertheless, the implementation of the latter process requires the separability of the log-likelihood function and this property has shown to be formally possible only under normality or, very recently, under expansions based on the Gaussian density (Del Brio et al., 2010).

In this paper we present a different approach to the joint distribution of financial asset returns that solves some of the aforementioned shortcomings, and produces reasonably good empirical results. Specifically, we propose a methodology to generalize the univariate SNP-type of distributions to a multivariate framework, which preserves the good performance of the SNP distributions in terms of generality and flexibility, since the marginal distributions of the proposed joint density are univariate SNP densities. The methodology may also incorporate the Gallant & Nychka's (1987) reformulations to ensure positivity, but the main stress is put on achieving the separability of the log-likelihood to be able to implement the two-stage estimation procedures.

2.1 The uncorrelated MGME

We start a our analysis by defining a naive case of the MGME distribution of uncorrelated variables (Definition 1) and afterwards we proceed to extend the definition to dependent variables through a simple linear transformation (Definition 2).

Definition 1. *The MGME pdf of a random vector* $x_t = (x_{1t}, x_{2t}, \ldots, x_{nt})' \in \mathbb{R}^n$ *of uncorrelated variables that uses a sequence of univariate pdfs,* $\{g_i(x_{it})\}_{i=1}^n$*, as basis, with finite non-central moments* $E[x_{it}^r] = \int x_{it}^r g_i(x_{it}) dx_{it} = \mu_{ir} < \infty$ $\forall i = 1, 2, \ldots, n$ *and* $\forall r \leq m$*, is defined as,*

$$F_k(\mathbf{x}_t; \boldsymbol{\gamma}) = \frac{1}{n} \left[\prod_{i=1}^{n} g_i(x_{it}) \right] \left[\sum_{i=1}^{n} w_{ik} \left(1 + \sum_{s=1}^{m} \gamma_{is}^k (x_{it}^s - \mu_{is})^k \right) \right], \ \forall k = 1, 2, \qquad (1)$$

where γ *is a matrix of parameters with general element* $\{\gamma_{si}\}$*,* w_{ik} *is the constant that make the marginal density of the* $i-th$ *dimension to integrate up to one, i.e.* $w_{i1} = 1$ *and*

$$w_{i2}^{-1} = \int \left[1 + \sum_{s=1}^{m} \gamma_{is}^2 (x_{it}^s - \mu_{is})^2 \right] g_i(x_{it}) dx_{it} = 1 + \sum_{s=1}^{m} \gamma_{is}^2 (\mu_{i,2s} - \mu_{is}^2), \ \forall i = 1, 2, ..., n. \quad (2)$$

It is noteworthy that the definition above encompasses many different multivariate distributions that can be obtained by assuming different distributions used as basis or by considering slightly different positivity transformations. Particularly the SNP family of distributions defined in Del Brio et al. (2010) are nested in the definition above by considering the standard Normal density as basis, $g_i(x_{it}) = \phi(x_{it}) = \frac{1}{\sqrt{2\pi}} e^{-x_{it}^2/2}$ $\forall i = 1, 2, ..., n$. The rationale behind this assessment lies in the fact that the Hermite polynomials (HP hereafter), those $H_{is}(x_{it})$ satisfying

$$\frac{d^s \phi(x_{it})}{dx_{it}^s} = (-1)^s \phi(x_{it}) H_s(x_{it}), \qquad (3)$$

also hold[1]

$$x_{it}^s - \mu_{is}^+ = \sum_{j=1}^{s/2-1} c_j H_{s-2j}(x_{it}), \tag{4}$$

$\mu_{is}^+ = \dfrac{s!}{2^{\frac{s}{2}}(s/2)!}$ being the $s-th$ order non-central moment of the standard Normal distribution and $\{c_j\}_{j=1}^{s}$ a particular parameter set. Hereafter we refer to the multivariate family of distributions represented in terms of the HP (i.e. Gram-Charlier series) as MGC (equation (5)).

$$MGC_k(\mathbf{x}_t; \mathbf{\Delta}) = \frac{1}{n}\left[\prod_{i=1}^{n} \phi\left(x_{it}\right)\right]\left[\sum_{i=1}^{n} \lambda_{ik}\left(1 + \sum_{s=1}^{m} \delta_{is}^k H_{is}(x_{it})^k\right)\right], \ \forall k = 1, 2, \tag{5}$$

where $\mathbf{\Delta}$ is a matrix of parameters with general element $\{\delta_{si}\}$ and λ_{ik} is the scaling constant such that

$$\lambda_{ik}^{-1} = \int \left[1 + \sum_{s=1}^{m} \delta_{is}^k H_{is}(x_{it})^k\right] \phi\left(x_{it}\right) dx_{it} = \begin{cases} 1, & \text{if } k = 1, \\ 1 + \sum\limits_{s=1}^{m} \delta_{is}^2 s!, & \text{if } k = 2, \quad \forall i = 1, 2, ..., n. \end{cases} \tag{6}$$

It is clear that when the standard Normal is used as basis and $k = 1$ the MGME renders the same density than the MGC, i.e. $F_1^+(\mathbf{x}_t; \boldsymbol{\gamma}) = GC_1(\mathbf{x}_t; \mathbf{\Delta})$, but this property does not hold for $k = 2$.[2] This fact deserves further discussion since is the basis of the idea of defining this new family of MGME distributions. On one hand the baseline case of $k = 1$ establishes that the MGME may be reformulated in terms of the GC (Type A) series and thus inheritates the good properties of such asymptotic expansion, particularly the capacity of approximating any probability density function (under certain regularity conditions). Nevertheless the representation in terms of the GME polynomials ($\{x_{it}^s - \mu_{is}^+\}_{s=1}^{m}$) is much more simple and easier to implement for empirical purposes, despite the fact that its polynomials do not form an orthogonal basis, unlike the HP polynomials. These simple polynomial structures

[1]Note that, without loss of generality, equation (4) shows the linear relation among Hermite and GME polynomials for s even. Analogously if s is odd $x_{it}^s - \mu_{is}^+$ follows a linear with the odd Hermite polynomials. Also note that these linear relations have no intercept. See Ñíguez & Perote (2007) for further details.

[2]Observe that the cross in $F_k^+(\mathbf{x}_t; \boldsymbol{\gamma})$ denotes that the Normal is used as basis of the MGME.

also permit to straightforwardly extend these expansions to other multivariate distributions beyond the Normal case.[3]

On the other hand the consideration of positivity restrictions (Jondeau & Rockinger, 2001) or positive transformations (Gallant & Nychka, 1987) is strictly required for obtaining a well-defined pdf, since positivity is not guaranteed when the expansions are truncated at the $m-th$ term. This problem is inherent of both GC and GME finite expansions but the latter involves straightforward positivity conditions, e.g. $0 \leq \gamma_{is} \leq \dfrac{1}{m\mu_{is}}$ $\forall s$ even, and $\gamma_{is} = 0$ $\forall s$ odd ($\forall i = 1, 2, ..., n$), and more simple expressions when transformations, as the one imposed for $k = 2$, are implemented. The positive transformation included in equation (1), however, is not the unique transformation to achieve positivity but we used it in definition 1 for the sake of simplicity.[4]

It is clear that the MGME distribution constitutes a well-defined pdf since is always positive and integrates up to one (Proof 1 in Appendix). Furthermore the statistical properties of the MGME distribution are easy to derive.

2.1.1 MGME properties

1. Marginal densities (Proof 2 in Appendix).

$$f_{ki}(x_{it}) = \begin{cases} g_i(x_{it}) \left[1 + \sum\limits_{s=1}^{m} \dfrac{\gamma_{is}}{n} \left(x_{it}^s - \mu_{is} \right) \right], & \text{if } k = 1, \\ g_i(x_{it}) \left[\dfrac{n-1}{n} + \dfrac{w_{i2}}{n} \left(1 + \sum\limits_{s=1}^{m} \gamma_{is}^2 \left(x_{it}^s - \mu_{is} \right)^2 \right) \right], & \text{if } k = 2. \end{cases} \tag{7}$$

2. Non-central moments (Proof 3 in Appendix).

$$E_k \left[x_{it}^r \right] = \begin{cases} \mu_{ir} + \sum\limits_{s=1}^{m} \frac{\gamma_{is}}{n} (\mu_{i,r+s} - \mu_{ir}\mu_{is}), & \text{if } k = 1, \\ \left[\frac{n-1}{n} + \frac{w_{ik}}{n} \right] \mu_{ir} + \frac{w_{ik}}{n} \sum\limits_{s=1}^{m} \gamma_{is}^2 \left[\mu_{i,2s+r} + \mu_{is}(\mu_{is}\mu_{ir} - 2\mu_{i,s+r}) \right], & \text{if } k = 2, \quad \forall r \in \mathbb{N}. \end{cases} \tag{8}$$

[3] Note that it is also feasible to obtain expansions of other distributions based on the derivatives of the distribution used as basis. Particularly for the Poisson, Gamma or Beta distributions they are the so-called Gram-Charlier Type B, Laguerre and Jacobi expansions respectively. Nevertheless these approaches are less tractable for empirical purposes than the one based on GME polynomials.

[4] León, Mencía & Sentana (2009) describe the parametric properties of univariate SNP distributions based on other alternative positive transformation.

3. Cumulative distribution function (Proof 4 in Appendix).

$$\Pr[x_1 \leq \overline{x}_1, \cdots, x_n \leq \overline{x}_n]_k = \frac{1}{n}\sum_{i=1}^{n} h_{ik}(\overline{x}_i)\left[\prod_{j=1,\ j\neq i}^{n}\int_{-\infty}^{\overline{x}_j} g_j\left(x_{jt}\right)dx_{jt}\right],\ \forall k=1,2. \tag{9}$$

where $h_{ik}(\overline{x}_i)$ stands for the cdf of the corresponding univariate GME distribution evaluated at $\overline{x}_i$ (see Ñíguez et al., 2007, for a closed formula of $h_{ik}(\overline{x}_i)$ for the Gaussian case).

2.2 The MGME with correlated variables

So far we have analysed the case of the MGME distribution of uncorrelated variables (Definition 1), but dependences among its marginals are incorporated to the MGME density in Definition 2 by considering a linear transformation of the type

$$\mathbf{u}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{x}_t = \mathbf{D}_t\mathbf{R}^{1/2}\mathbf{x}_t, \tag{10}$$

where the (positive definite) variance and covariance matrix, $\mathbf{\Sigma}_t = \mathbf{\Sigma}_t^{1/2}\mathbf{\Sigma}_t^{1/2} = \mathbf{D}_t\mathbf{R}\mathbf{D}_t = \mathbf{D}_t\mathbf{R}^{1/2}\mathbf{R}^{1/2}\mathbf{D}_t$, has been decomposed in the diagonal matrix of conditional deviations, $\mathbf{D}_t = diag\{\sigma_{1t}, \cdots, \sigma_{nt}\}$, and the correlation matrix, $\mathbf{R}$.

Definition 2. *The MGME pdf of a random vector $u_t = (u_{1t}, u_{2t}, \ldots, u_{nt})' \in \mathbb{R}^n$ and using the multivariate pdf $G(u_t; \Sigma_t, \theta)$ (with mean 0, conditional variance and covariance matrix Σ_t, density parameters θ, marginal densities $g_i(x_{it})$ and non-central moments denoted by μ_r, $\forall i = 1, 2, ..., n$) as basis, is defined as,*

$$F_k(\mathbf{u}_t; \boldsymbol{\gamma}, \mathbf{\Sigma}_t, \boldsymbol{\theta}) = \frac{1}{n}G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta})\sum_{i=1}^{n} w_{ik}\left[1 + \sum_{s=1}^{m}\gamma_{is}^{k}\left(x_{it}^{s} - \mu_{is}\right)^k\right],\ \forall k=1,2, \tag{11}$$

where x_{it} is the corresponding component of the inverse transformation in equation (10), e.g. for the bivariate case and the symmetric eigenvalue decomposition of the variance and covariance matrix

$$x_{it} = \frac{1}{2}\sum_{j=1}^{2}\left(\frac{1}{\sqrt{1+\rho}} + (-1)^{i+j}\frac{1}{\sqrt{1-\rho}}\right)\frac{u_{jt}}{\sigma_{jt}},\ \forall i=1,2, \tag{12}$$

where σ_{jt} captures the conditional deviation of u_{jt} ($\forall j = 1, 2$) and $|\rho| < 1$ the correlation between variables u_{1t} and u_{2t}.[5]

Special interest deserves the case where the multivariate density used as basis is elliptical and thus may be expressed as

$$G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta}) = |\mathbf{\Sigma}_t|^{-1/2} \varphi_n \left(\mathbf{u}_t' \mathbf{\Sigma}_t^{-1} \mathbf{u}_t\right), \tag{13}$$

where $\varphi_n(z)$, $z \geq 0$, is some generating function such that

$$\int_0^\infty z^{n/2-1} \varphi_n(z)\, dz = \frac{\Gamma(\frac{n}{2})}{\pi^{n/2}}. \tag{14}$$

In particular for $\varphi_n(z) = \frac{e^{-z/2}}{(2\pi)^{n/2}}$ we obtained the MGME that expands the multivariate Normal (MN hereafter) (equation (15)) and for $\varphi_n(z) = \left(1 + \frac{z}{\nu}\right)^{-(\nu+n)/2}$, ν being the degrees of freedom, the expansion of the multivariate Student's t (MST hereafter) (equation (16)).[6]

$$\begin{aligned} F_k^+(\mathbf{u}_t; \boldsymbol{\gamma}, \mathbf{\Sigma}_t) &= \frac{1}{(2\pi)^{\frac{n}{2}} n} |\mathbf{\Sigma}_t|^{-\frac{1}{2}} \exp\left\{-\frac{\mathbf{u}_t' \mathbf{\Sigma}_t^{-1} \mathbf{u}_t}{2}\right\} \sum_{i=1}^n w_{ik} \left[1 + \sum_{s=1}^m \gamma_{is}^k \left(x_{it}^s - \mu_{is}^+\right)^k\right] & (15) \\ F_k^*(\mathbf{u}_t; \boldsymbol{\gamma}, \mathbf{\Sigma}_t, \nu) &= \frac{\Gamma\left(\frac{\nu+n}{2}\right) |\mathbf{\Sigma}_t|^{-\frac{1}{2}} \left[1 + \frac{\mathbf{u}_t' \mathbf{\Sigma}_t^{-1} \mathbf{u}_t}{(\nu-2)}\right]^{-\frac{\nu+n}{2}} \sum_{i=1}^n w_{ik} \left[1 + \sum_{s=1}^m \gamma_{is}^k \left(x_{it}^s - \mu_{is}^*\right)^k\right]}{n(\pi(\nu-2))^{\frac{n}{2}} \Gamma\left(\frac{\nu}{2}\right)} & (16) \\ \forall k &= 1, 2, \end{aligned}$$

μ_{is}^+ and $\mu_{is}^* = \mu_{is}^+ \frac{(\nu-2)^{s/2-1}}{(\nu-s)(\nu-s-2)(\nu-s-4)\cdots(\nu-4)}$ $\forall s$ even (0 $\forall s$ odd) being the $s-th$ order non-central moment of the standard Normal and standard Student's t distribution, respectively. Note that the MN and the MST are nested in (15) and (16), respectively, and that $F_k^*(\mathbf{u}_t; \boldsymbol{\gamma}, \mathbf{\Sigma}_t, \nu)$ tends to $F_k^+(\mathbf{u}_t; \boldsymbol{\gamma}, \mathbf{\Sigma}_t)$ as ν goes to infinity.

The statistical properties of the "Non-standard" MGME family of distributions in Definition 2 are easily worked out by taking into account the properties of the density

[5] It must be noted that the decomposition of the variance and covariance matrix is not unique. Particularly, the non-symmetric eigenvalue or the Cholesky decomposition may also be implemented. Even the trivial transformation for uncorrelated variables (i.e. $x_{it} = \frac{u_{it}}{\sigma_{it}}$ $\forall i = 1, 2, ..., n$) might be used in the polynomial structure, since correlations have been taken into account in the multivariante distribution used as basis, $G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta})$.

[6] There exist many other specifications for the multivariate Student's t. We opt for the more standard case (equation (1.1) in Kotz & Nadarajah, 2004).

used as basis, the transformation (10) and the standard case in equations (7), (8) and (9). For example, the MGME has not $\mathbf{0}$ mean unless $\gamma_{si} = 0$ $\forall s$ odd and $\forall i$, the variance and covariance matrix is proportional to $\mathbf{\Sigma}_t$, and all the moment and co-moment structure (including skewness and kurtosis) depend on the matrix of parameters $\boldsymbol{\gamma}$. More interestingly, the MGME distributions that use the MN as basis preserve the "separability" property introduced in Engle (2002) and Engle and Sheppard (2001), since the log-likelihood function can be split in the log-likelihood of the volatility part, $L_V(\mathbf{u}_t, \boldsymbol{\alpha})$ (equation (17)), and the log-likelihood of the MGME in terms of the standardized variables ($\boldsymbol{\varepsilon}_t = \mathbf{D}_t^{-1}\mathbf{u}_t$), $L_{GME_k}(\boldsymbol{\varepsilon}_t, \boldsymbol{\alpha}, \boldsymbol{\rho})$ (equation (18)), where $\boldsymbol{\alpha}$, $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ stand for the matrices containing the parameters of the conditional variances, correlations and expansion terms of the MGME distribution, respectively (Proof 5 in Appendix).

$$\begin{aligned} L_V(\mathbf{u}_t, \boldsymbol{\alpha}) &= -\frac{1}{2}\sum_{i=1}^{n}\left[T\log(2\pi) + \sum_{t=1}^{T}\left(\ln(\sigma_{it}^2) + \frac{u_{it}^2}{\sigma_{it}^2}\right)\right] \\ &= -\frac{1}{2}\sum_{i=1}^{n}\left[T\log(2\pi) + L_{V_i}(\boldsymbol{\alpha}_i)\right], \end{aligned} \tag{17}$$

$$\begin{aligned} L_{GME_k^+}(\boldsymbol{\varepsilon}_t, \boldsymbol{\alpha}, \boldsymbol{\rho}, \boldsymbol{\gamma}) &= -\tfrac{1}{2}\sum_{t=1}^{T}\left\{\ln|\mathbf{R}| + \boldsymbol{\varepsilon}_t'\mathbf{R}^{-1}\boldsymbol{\varepsilon}_t - 2\ln\left[\sum_{i=1}^{n} w_{ik}\left(1 + \sum_{s=1}^{m}\gamma_{is}^k\left(x_{it}^s - \mu_{is}^+\right)^k\right)\right]\right\} \\ \forall k &= 1, 2. \end{aligned} \tag{18}$$

The latter property allows to implement two-step estimation methods since the parameters of the conditional variances can be consistently estimated by independent QML estimation under Gaussian hypotheses (first step) and correlations can be estimated jointly with the rest of the MGME parameters by LIML applied to the log-likelihood concentrated on the estimates of the conditional variances, $L_{GME_k}(\boldsymbol{\varepsilon}_t, \widehat{\boldsymbol{\alpha}}, \boldsymbol{\rho}, \boldsymbol{\gamma})$ where $\widehat{\boldsymbol{\alpha}} = \arg\max\{L_{MV}(\mathbf{u}_t, \boldsymbol{\alpha})\}$ (second step).[7] Such a procedure simplifies by far the estimation compared to the jointly estimation that is theoretically the only valid methodology if the assumed distribution is not Normal (i.e. for financial data applications). Del Brio et al. (2010) extended this two-step methodology to non-Gaussian distributions of the type displayed in equation (5). These

[7] Three-step estimation methods can be also implemented analogously, the first step estimating the conditional mean parameters and conditioning on these estimates the second and third steps.

authors argued that as the GC is a valid asymptotic expansion the second step must also be necessarily consistent even when the true distribution is not MGC. The same result holds for our MGME approach, since it encompasses the MGC distribution.

3 Application to exchange rates portfolio

3.1 The model

Let us consider a portfolio of n assets and let $\mathbf{r}_t$ the vector containing the returns on these assets at time t. We assume that the distribution of $\mathbf{r}_t$ conditioned on the information set $\mathbf{\Omega}_{t-1}$ belongs to the MGME family with conditional mean, $\mathbf{E}(\mathbf{r}_t|\mathbf{\Omega}_{t-1}) = \boldsymbol{\mu}_t(\boldsymbol{\phi})$, and and conditional variance and covariance matrix $\mathbf{E}(\mathbf{u}_t\mathbf{u}_t'|\mathbf{\Omega}_{t-1}) = \mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho})$ (equations (19), (20) and (21)), where $\boldsymbol{\phi}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ are the matrices including conditional mean, variance, correlation and MGME expansion parameters, respectively. We model conditional mean and variance for every variable as AR(1)-GARCH(1,1) (equations (22) and (23), $\circ$ being the Hadamard product computed by element by element multiplication) and we assume the CCC model (Bollerslev, 1990) for modeling correlations, which guarantees a positive definite variance and covariance matrix.[8]

$$
\begin{aligned}
\mathbf{r}_t &= \boldsymbol{\mu}_t(\boldsymbol{\phi}) + \mathbf{u}_t, && (19)\\
\mathbf{u}_t|\Omega_{t-1} &\sim MGME_k(\mathbf{0}, \mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho}), \boldsymbol{\gamma}), && (20)\\
\mathbf{\Sigma}_t(\boldsymbol{\alpha},\boldsymbol{\rho}) &= \mathbf{D}_t(\boldsymbol{\alpha})\mathbf{R}(\boldsymbol{\rho})\mathbf{D}_t(\boldsymbol{\alpha}), && (21)\\
\boldsymbol{\mu}_t(\boldsymbol{\phi}) &= \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1'\mathbf{r}_{t-1} && (22)\\
\mathbf{D}_t^2 &= diag\{\alpha_{i0}\} + diag\{\alpha_{i1}\} \circ \mathbf{u}_{t-1}\mathbf{u}_{t-1}' + diag\{\alpha_{i2}\} \circ \mathbf{D}_{t-1}^2, && (23)
\end{aligned}
$$

We consider different SNP models that are nested in the MGME family of distributions. In particular we consider four different expansions using the MN as basis: the MGC_k (equation (5)) and the $MGME_k^+$ (equation (15)) $\forall k = 1, 2$. We choose these specifications to show that

[8] We choose the more benchmarked model for first and second moments since we focus on showing the advantages of the MGME distribution compared to other multivariate SNP alternatives. Of course our model admits more complex multivarite volatilitiy structures, particularly the DCC model by Engle (2002).

MGC_1 is nested in the the MGME family since its density can be obtained as $MGME_1^+$ and that the "positive" $MGME_2^+$ may be a more simple and accurate formulation than the MGC_2. Furthermore, to show the potential advantages of the MGME when using other densities as basis we also include a version of the MGME using the MST, $MGME_1^*$ (equation (16)), which we compare with the MST as benchmark. We expand all these densities to the eighth term although some of the parameters are constrained to zero, after testing their non-significance (e.g. the odd parameters). For the sake of simplicity in the empirical application presented below we consider a bivariate portfolio, the applications of MGME densities using the MN as basis to larger portfolios, however, are not computationally demanding since they may be consistently estimated in two (or three) steps.

3.2 The Data

In this section we investigate the empirical performance of the six bivariate models discussed above: MGC_1, MGC_2, $MGME_1^+$, $MGME_2^+$, $MGME_1^*$ and MST. The data used are (daily) percent log returns computed as $r_{it} = 100 \log (P_{it}/P_{it-1})$ from a observed sample $\{P_{it}\}_{t=1}^T$ of (daily) FX of different currencies with respect to the British Pound. In particular we analyse the bivariate distribution of the US Dolar/British Pound ($/£) and EU Euro/British Pound (€/£) FX and the Chinese Yuan/British Pound (Y/£) and the Japanese Yen/British Pound (y/£) FX. All series are sampled over the period June 20, 1995 to October 2, 2009 for a total of 3,560 observations. The data were obtained from Datastream. Table 1 reports some descriptive statistics for the total samples. The unconditional distribution of any of these series shows clearly non-Gaussian features, such as (mild) skewness, and a remarked excess of kurtosis over the Normal distribution, Japanese y/£ FX being by far the more leptokurtic and extreme valued series. Regarding the correlation coefficients, the $/£ and the €/£ FX are higher correlated than the Asian currencies between them (Y/£ and y/£ FX).

[Insert Table 1]

3.3 Estimation and empirical analysis

The estimation of the MGME models that expands the MN is carried out in two stages by (Q)MLE techniques. In the first stage, the parameters of the AR(1)-GARCH(1,1) model are estimated (under normality) independently for each asset and, in the second stage, the centered and standardized residuals from the previous step are used to estimate correlations and the rest of the density parameters. The likelihood functions in each step are maximized using the Berndt, Hall, Hall & Hausman (1974) algorithm. The first stage yields QMLE which is consistent and asymptotically normal, although not efficient. In cases of density misspecification the second stage MLE is not a priory consistent, since we use a truncated SNP density, but it may be more efficient than the second step QMLE (i.e. under the MN distribution). However, we argue that the better our truncated SNP density can approximate the "true" distribution, the more efficient is our second stage MLE. Finally, following Pagan (1996) a further Newton-Raphson iteration without line search for the joint model is performed from the first and second stage (Q)MLE: the estimates do not change but the information matrix is now block diagonal, thus obtaining estimators asymptotically equivalent to joint QMLE. Robust standard errors were computed following Bollerslev & Wooldridge (1992). The models that include the MST are jointly estimated by MLE since two step procedures cannot be theoretically applied because the log-likelihood is not separable. Nevertheless when multi-step procedures are "ad hoc" implemented to the MST the estimation results do not differ significantly from the one-step approach (Bauwens & Laurent, 2005, and Jondeau & Rockinger, 2005). We observe that the estimation of the MGME models is not computationally very demanding providing that starting values are chosen properly. As it is known that MGME densities may present multiple local modes, the optimization is monitored using different starting values to ensure that the obtained (Q)MLE estimates are global optima.

Table 2 presents the estimation of the parameters of the aforementioned bivariate models, which follow the same notation used in previous sections with the exception of the parameters of the expansion terms of the MGC distribution that, for the sake of simplicity, have the same notation than the parameters of MGME expansions. Particularly, ϕ_{is} $(s = 0, 1)$ stand for the AR(1) parameters of the conditional means, α_{is} ($s = 0, 1, 2$) for the GARCH(1,1) parameters of the conditional variances, ρ for the unconditional correlation parameter, γ_{is} ($s = 2, 4, 6, 8$) for the $s - th$ order polynomial weighting parameter of the expansions and ν for the degrees of freedom (df hereafter) of the MST. T-ratios for robust standard errors are in parenthesis next to the parameter estimates and an asterisk signals the insignificant parameters (at 5 percent confidence level).

Regarding the specification of the MGC and MGME models we considered expansions truncated at the eighth term, but systematically non-significant parameters in all distributions were removed. Specifically, the estimated densities for both portfolios are unconditionally symmetric since the odd parameters, γ_{is} $s = 3, 5, 7$ ($i = 1, 2$) are not significant at any reasonable significance level. Analogously parameters γ_{i4} and γ_{i6} ($i = 1, 2$) of the positive distributions (MGC_2 and $MGME_2^+$) and parameters γ_{i6} and γ_{i8} ($i = 1, 2$) of the expansion of the MST ($MGME_1^*$) are also omitted. The selection of the parametric structure for these expansions was selected according to linear restriction (Wald) tests but there exists a clear intuition behind these particular parametric structures, since the parameters of the MGME expansions capture the weights assigned to the deviations of the moments of the empirical distribution from the corresponding moments of the distribution used as basis. Therefore it seems that deviations around the mean (γ_{i2}) and the deviations for extreme values (γ_{i8}) have significant role for the expansion of the MN but the most relevant parameters are γ_{i2} and γ_{14} when the MST is used as basis, since df parameter (ν) may capture the thick tailed behavior (i.e., a larger expansion is not required).

The models are compared according to accuracy criteria and for this reason the log-likelihood value (LnL) and the Schwarz Bayesian Information Criterion (BIC) are displayed in the last two rows of Table 2.[9] According to these criteria we observe the following evidence: (1) The MGME models provide a notably better goodness-of-fit than the distributions used as basis by themselves (Gaussian[10] or Student's t). (2) If positivity transformations are not

[9] We choose this statistic instead of the Akaike since the BIC, defined as $BIC = -lnL + pln(T)/2$ (p being stands for the number of the parameters of the model), has optimal properties (Geweke & Meese 1981).

[10] The log-likelihood and BIC values of the benchmarked MN are not reported in Table 1. These values are -2371.67 and 2400.25 for the joint distribution of \$/£ and €/£ FX, and -4381.69 and 4410.27 for the joint distribution of Y/£ and y/£ FX, respectively.

implemented the MGC and GME yield the same density, although the latter specification seems to increase parameters significance (it is the additional information of the empirical moments from those of the density used as basis what significantly explains the expansion terms). (3) The densities that do not incorporate positivity transformations seem to be more accurate since those transformations imply certain restrictions in the parametric space, nevertheless such transformations may be strictly necessary in different applications which require estimating recursively the density (e.g. density forecasting). (4) The comparison of the accuracy obtained expanding MST instead of MN models is misleading because both models are not nested, however it seems that the expansion of the MN requires additional terms to capture fat tails compared to the MST that also has the df parameter for such purposes (actually df parameter increases when the MST is expanded because the terms of the expansion capture some part of the extreme values). On the other hand the expansion of the MST (or other non-normal distribution) has not been proved to yield consistent estimates if two-step procedures are implemented (which may be an important drawback for large portfolios).

We also observe the usual small structure in the conditional mean, high persistence in the conditional variance and the correlation parameters capture the fact that the \$/£ and the €/£ FX are higher correlated than the Y/£ and y/£ FX (although it must be noted that parameter ρ does not accounts exactly for the correlation coefficient since variances and covariances depend also on the parameters of the expansion terms).

[Insert Table 2]

Finally, we include a picture of the bivariate GME distribution of the Y/£ and y/£ FX in Figure 1. This plot illustrates the type of distributions obtained by using the type of expansions proposed in this article and how they can capture the thick tails featured by financial returns. We also represent in Figures 2A and 2B the the marginal fitted densities of the y/£ FX distribution (the most leptokurtic of the analysed series) according to different specifications compared to the empirical distribution (histogram). These plots reinforce the evidence commented in points (1) and (3) above, since the GME densities clearly outperform the Normal and the Student's t (specially at the tails) and the "non-positive" GME (GME) presents a more accurate data fit than the "positive" GME (GME2).

[Insert Figures 1 and 2]

4 Concluding Remarks

The literature on the multivariate volatility modeling of portfolio returns has traditionally focused on the time-varying first and second conditional moments of the asset returns distribution. Nevertheless the abundant evidence of leptokurtosis in portfolio distribution raises the need of modeling higher order moments by using more flexible specifications than the traditional but unreliable Gaussian assumption. For this purpose there have being proposed the use of several parametric distributions or copulas, but these alternatives are either not flexible enough to incorporate salient empirical regularities of financial returns (such as heavy tails, possible multimodality, skewness, etc.) or analytically intractable if they incorporate those features. On the other hand the SNP approach based on MGC expansions allows to straightforwardly address these topics since, not only because they can fit any target density through their general an flexible parametric structure, but also they present an analytical specification that is tractable due to the orthogonal structure of Hermite polynomials. In this paper we propose an alternative SNP family of distributions, the MGME, that are much simpler (the polynomials of the expansion do not require orthogonality) and thus easier to implement even when using as basis any distribution with the only requirement of having as many finite moments as the expansion length for every dimension. Furthermore our approach generalizes the univariate GME proposed by Ñíguez & Perote (2007) to a multivariate framework.

If Gaussian density is used as basis and positive transformations are not implemented the MGME is just a respecification of the MGC but if positivity is imposed the MGME yields a more simple and accurate formulation. In these cases the MGME also admits the decomposition of the likelihood function proposed in Engle (2002), which permits to estimate independently the volatility processes of every asset and, in a second stage, the rest of the density parameters (correlations and the parameters of the expansion terms), thus solving the "dimensionality curse" of multivariate volatility models .

We compare the empirical performance of different types of MGME for modeling the distribution of FX portfolio returns (\$/£-€/£ and Y/£-y/£ FX), in relation to the MN and MST, taken as basis, and alternative MGC distributions. The results show that the MGME specifications outperform not only the MN and MST but also might be superior than the MGC distributions when positive transformations are implemented, and thus the MGME may be an interesting an easy-to-implement SNP tool for modeling portfolio distributions.

Appendix

This appendix includes the proofs of the properties of the multivariate GME densities presented in Section 2. *Proof 1* shows that multivariate SNP densities integrate up to one; *Proof 2, Proof 3 and Proof 4* provide closed forms for marginal distributions, moments and cdfs, respectively; *Proof 5* shows the separability of the log-likelihood for the MGME.

Proof 1.

$$\begin{aligned}
&\int \cdots \int F_k(\mathbf{x}_t) dx_{1t} \cdots dx_{nt} \\
=\ & \frac{1}{n} \int \cdots \int \left\{ \prod_{i=1}^{n} g_i\left(x_{it}\right) \right\} \left\{ \sum_{i=1}^{n} w_{ik} \left[1 + \sum_{s=1}^{m} \gamma_{is}^{k} \left(x_{it}^{s} - \mu_{is}\right)^{k} \right] \right\} dx_{1t} \cdots dx_{nt} \\
=\ & \frac{1}{n} \sum_{i=1}^{n} \left[w_{ik} \int \left(1 + \sum_{s=1}^{m} \gamma_{is}^{k} \left(x_{it}^{s} - \mu_{is}\right)^{k} \right) g_i\left(x_{it}\right) dx_{it} \prod_{j=1,\ j \neq i}^{n} \int g_j\left(x_{jt}\right) dx_{jt} \right] \\
=\ & \frac{1}{n} n = 1,\ \forall k = 1, 2 \ \blacksquare
\end{aligned}$$

Proof 2.

$$\begin{aligned}
f_{ik}(x_{it}) &= \int \cdots \int F_k(\mathbf{x}_t) dx_{1t} \cdots dx_{i-1,t} dx_{i+1,t} \cdots dx_{nt} \\
&= \frac{1}{n} g_i(x_{it}) w_{ik} \left(1 + \sum_{s=1}^{m} \gamma_{is}^{k} \left(x_{it}^{s} - \mu_{is}\right)^{k} \right) \prod_{j=1,\ j \neq i}^{n} \int g_j\left(x_{jt}\right) dx_{jt} \\
&+ \frac{1}{n} g_i(x_{it}) \sum_{j=1,\ j \neq i}^{n} \left[\prod_{l=1,\ l \neq i}^{n} w_{ik} \int g_l\left(x_{lt}\right) \left(1 + \sum_{s=1}^{m} \gamma_{ls}^{k} \left(x_{lt}^{s} - \mu_{ls}\right)^{k} \right) dx_{lt} \right] \\
&= \frac{1}{n} g_i(x_{it}) w_{ik} \left(1 + \sum_{s=1}^{m} \gamma_{is}^{k} \left(x_{it}^{s} - \mu_{is}\right)^{k} \right) + \frac{n-1}{n} g_i(x_{it}), \quad \forall k = 1, 2 \textit{ and } \forall i = 1, 2, ..., n \ \blacksquare
\end{aligned}$$

Furthermore, If $k = 1$ *(i.e.* $w_{i1} = 1$*) then*

$$f_{i1}(x_{it}) = g(x_{it})\left[(n-1) + \left(1 + \sum_{s=1}^{m}\gamma_{is}\left(x_{it}^{s} - \mu_{is}\right)\right)\right]\frac{1}{n} = g_i(x_{it})\left[1 + \sum_{s=1}^{m}\frac{\gamma_{is}}{n}\left(x_{it}^{s} - \mu_{is}\right)\right] \blacksquare$$

Proof 3.

$$\begin{aligned} E_1\left[x_{it}^r\right] &= \int x_{it}^r f_{i1}(x_{it})dx_{it} = \int x_{it}^r g_i\left(x_{it}\right)dx_{it} + \sum_{s=1}^{m}\frac{\gamma_{is}}{n}\int x_{it}^r\left(x_{it}^s - \mu_{is}\right)g_i(x_{it})dx_{it} \\ &= \mu_{ir} + \sum_{s=1}^{n}\frac{\gamma_{si}}{n}(\mu_{i,r+s} - \mu_{ir}\mu_{is}) \blacksquare \end{aligned}$$

$$\begin{aligned} E_2\left[x_{it}^r\right] &= \int x_{it}^r f_{i2}(x_{it})dx_{it} \\ &= \frac{n-1}{n}\int x_{it}^r g_i(x_{it})dx_{it} + \frac{1}{n}w_{ik}\int x_{it}^r g_i(x_{it})dx_{it} + \frac{1}{n}\sum_{s=2}^{m}\gamma_{is}^2\int x_{it}^r\left(x_{it}^s - \mu_{is}\right)^2 g_i(x_{it})dx_{it} \\ &= \left[\frac{n-1}{n} + \frac{w_{ik}}{n}\right]\mu_{ir} + \\ &\quad \frac{w_{ik}}{n}\sum_{s=1}^{m}\gamma_{is}^2\left(\int x_{it}^{2s+r}g_i(x_{it})dx_{it} + \mu_{is}^2\int x_{it}^r g_i(x_{it})dx_{it} - 2\mu_{is}\int x^{s+r}g_i(x_{it})dx_{it}\right) \\ &= \left[\frac{n-1}{n} + \frac{w_{ik}}{n}\right]\mu_{ir} + \frac{w_{ik}}{n}\sum_{s=1}^{m}\gamma_{is}^2\left[\mu_{i,2s+i} + \mu_{is}(\mu_{is}\mu_{ir} - 2\mu_{i,s+r})\right], \end{aligned}$$

$\forall r \in \mathbb{N}$ and provided that $g_i(x_{it})$ has finite moments at least up to the *order* $2s + r$.

Proof 4.

$$\begin{aligned} &\Pr[x_1 \le \overline{x}_1, \cdots, x_n \le \overline{x}_n]_k \\ &= \frac{1}{n}\int_{-\infty}^{\overline{x}_1}\cdots\int_{-\infty}^{\overline{x}_n}\left[\prod_{i=1}^{n}g_i\left(x_{it}\right)\right]\left\{\sum_{i=1}^{n}w_{ik}\left(1 + \sum_{s=1}^{m}\gamma_{is}^k\left(x_{it}^s - \mu_{is}\right)^k\right)\right\}dx_{1t}\cdots dx_{nt} \\ &= \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{\overline{x}_i}w_{ik}\left(1 + \sum_{s=1}^{m}\gamma_{is}^k\left(x_{it}^s - \mu_{is}\right)^k\right)g_i\left(x_{it}\right)dx_{it}\prod_{j=1,\ j\neq i}^{n}\int_{-\infty}^{\overline{x}_j}g_j\left(x_{jt}\right)dx_{jt} \\ &= \frac{1}{n}\sum_{i=1}^{n}h_{ik}(\overline{x}_i)\prod_{j=1,\ j\neq i}^{n}\int_{-\infty}^{\overline{x}_j}g_j\left(x_{jt}\right)dx_{jt}, \quad \forall k = 1, 2 \blacksquare \end{aligned}$$

Proof 5.

$$
\begin{aligned}
L_{GME_k^+}(\mathbf{u}_t,\boldsymbol{\alpha},\boldsymbol{\rho},\boldsymbol{\gamma}) &= -\frac{1}{2}\sum_{t=1}^{T}\big(n\ln(2\pi)+\ln|\boldsymbol{\Sigma}_t|+\mathbf{u}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{u}_t\big) \\
&\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_{ik}\left(1+\sum_{s=1}^{m}\gamma_{is}^{k}\left(x_{it}^{s}-\mu_{is}^{+}\right)^{k}\right)\right]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\big(n\ln(2\pi)+\ln|\mathbf{D}_t\mathbf{R}\mathbf{D}_t|+\mathbf{u}_t'\mathbf{D}_t^{-1}\mathbf{R}^{-1}\mathbf{D}_t^{-1}\mathbf{u}_t\big) \\
&\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_{ik}\left(1+\sum_{s=1}^{m}\gamma_{is}^{k}\left(x_{it}^{s}-\mu_{is}^{+}\right)^{k}\right)\right]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\big(n\ln(2\pi)+2\ln|\mathbf{D}_t|+\ln|\mathbf{R}|+\boldsymbol{\varepsilon}_t'\mathbf{R}^{-1}\boldsymbol{\varepsilon}_t\big) \\
&\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_{ik}\left(1+\sum_{s=1}^{m}\gamma_{is}^{k}\left(x_{it}^{s}-\mu_{is}^{+}\right)^{k}\right)\right]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\big(n\ln(2\pi)+2\ln|\mathbf{D}_t|+\mathbf{u}_t'\mathbf{D}_t^{-2}\mathbf{u}_t-\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t+\ln|\mathbf{R}|+\boldsymbol{\varepsilon}_t'\mathbf{R}^{-1}\boldsymbol{\varepsilon}_t\big) \\
&\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_{ik}\left(1+\sum_{s=1}^{m}\gamma_{is}^{k}\left(x_{it}^{s}-\mu_{is}^{+}\right)^{k}\right)\right]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{n}\left[\ln\left(2\pi\sigma_{it}^{2}\right)+\frac{\mathbf{u}_t^{2}}{\sigma_{it}^{2}}\right]-\frac{1}{2}\sum_{t=1}^{T}\left(\ln|\mathbf{R}|+\boldsymbol{\varepsilon}_t'\mathbf{R}^{-1}\boldsymbol{\varepsilon}_t\right) \\
&\quad +\sum_{t=1}^{T}\ln\left[\sum_{i=1}^{n}w_{ik}\left(1+\sum_{s=1}^{m}\gamma_{is}^{k}\left(x_{it}^{s}-\mu_{is}^{+}\right)^{k}\right)\right]+\frac{1}{2}\sum_{t=1}^{n}\boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t-T\ln(n) \\
&= L_V(\mathbf{u}_t,\boldsymbol{\alpha})+L_{GME_k^+}(\boldsymbol{\varepsilon}_t,\boldsymbol{\alpha},\boldsymbol{\rho},\boldsymbol{\gamma})+\kappa,\ \forall k=1,2\blacksquare
\end{aligned}
$$

References

[1] Alexander, C. O. & Chibumba, A. M. (1996). Multivariate orthogonal factor GARCH. University of Sussex Discussion Papers in Mathematics.

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

[3] Bauwens, L. & Laurent, S. (2005). A New Class of Multivariate Skew Densities, with Application to Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics*, 23, 346-354.

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

[5] Berndt, E. K., Hall, B. H., Hall, R. E. & Hausman, J. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economics and Social Measurement,* 4, 653-665.

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

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

[8] Chen, X., Fan, Y. & Tsyrennikov, V. (2006). Efficient Estimation of Semiparametric Multivariate Copula Models. *Journal of the American Statistical Association*, 101, 1228-1240.

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

[10] Del Brio, E. B., Ñíguez, T. M. & Perote, J. (2010). Multivariate Semi-Nonparametric Distributions with Dynamic Conditional Correlations. *International Journal of Forecasting*, forthcoming.

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

[12] Engle, R. F. & Kelly, B. (2007). Dynamic Equicorrelation. Proceedings of the 27th International Symposium on Forecasting, New York.

[13] Engle, R. F. & Kroner, K. F. (1995). Multivariate Simultaneous Generalized ARCH. *Econometric Theory*, 11, 122-150.

[14] Engle, R. F., Ng, V. K. & Rothschild, K. F. (1990). Asset Pricing with Factor-ARCH Covariance Structure: Empirical Estimates for Treasure Bills. *Journal of Econometrics*, 45, 213-237.

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

[16] Gallant, R. & Nychka, D. (1987). Seminonparametric Maximum Likelihood Estimation. *Econometrica*, 55, 363-390.

[17] Gallant, R. & Tauchen, G. (1989). Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications. *Econometrica*, 57, 1091-1120.

[18] Geweke, J. F. & Meese, R. (1981). Estimating Regression Models of Finite but Unknown Order. *International Economic Review*, 22, 55-70.

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

[20] Jondeau, E. & Rockinger, M. (2005). Conditional Asset Allocation under Non-Normality: How Costly Is the Mean-Variance Criterion? *EFA 2005 Moscow Meetings Discussion Paper.*

[21] Kotz, S. & Nadarajah, S. (2004). *Multivariate T-Distributions and Their Applications.* Cambridge University Press: Cambridge.

[22] León, A., Mencía, J. & Sentana, E. (2009). Parametric Properties of Semi-Nonparametric Distributions, with Applications to Option Valuation. *Journal of Business and Economic Statistics* 27, 176-192.

[23] Ñíguez, T. M. & Perote, J. (2007). Semi-Parametric Density Expansions: Orthogonality vs Simplicity. Proceedings of the 27th Symposium on Forecasting, New York.

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

[25] Pagan, A. R. (1986). Two Stage and Related Estimators and their Applications. *Review of Economic Studies*, 53, 517-538.

[26] Polanski, A. & Stoja, E. (2010). Dynamic Density Forecasts for Multivariate Asset Returns. Mimeo.

[27] Sentana, E. & Fiorentini, G. (2001). Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models. *Journal of Econometrics*, 102, 143-164.

[28] Van der Weide, R. (2002). GO-GARCH: A Multivariate Genaralized Orthogonal GARCH Model. *Journal of Applied Econometrics*, 17, 549-564.

Tables and Figures

TABLE 1

Daily percent log returns descriptive statistics

	Portfolio FX \$/£ - FX €/£		Portfolio FX Y/£ - FX y/£	
	FX \$/£	FX €/£	FX Y/£	FX y/£
Sample	20/06/1995 - 2/10/2009			
Observations	3560			
Mean	-0.0024637	-0.0020004	-0.0082610	-0.000097564
Maximum	4.47445	2.70093	3.26910	8.27608
Minimum	-3.91829	-3.14019	-3.95439	-6.23441
St. Dev.	0.53845	0.47118	0.54747	0.81205
Skewness	-0.17983	-0.30995	-0.22340	-0.40868
Kurtosis	5.06471	3.37101	3.92932	10.56811
Correlation	0.31747		0.075097	

TABLE 2

Estimation results

Mean equation: $r_{it} = \phi_{i0} + \phi_{i1} r_{i,t-1} + u_{it},\ \ u_{it} = \varepsilon_{it}\sigma_{it},$

Variance equation: $\sigma^2_{it} = \alpha_{i0} + \alpha_{i1} u^2_{i,t-1} + \alpha_{i2}\sigma^2_{i,t-1},\ \ i = 1, 2.$

MGC equation: $F_k(\mathbf{u}_t; \boldsymbol{\gamma}, \boldsymbol{\Sigma}_t) = \frac{1}{2} G(\mathbf{u}_t; \boldsymbol{\Sigma}_t) \left[\sum_{i=1}^{2} \lambda_{ik} \left(1 + \sum_{s=1}^{8} \gamma^k_{is} H_{is}(x_{it})^k \right) \right],\ k = 1, 2$

MGME equation: $F_k(\mathbf{u}_t; \boldsymbol{\gamma}, \boldsymbol{\Sigma}_t, \nu) = \frac{1}{2} G(\mathbf{u}_t; \boldsymbol{\Sigma}_t, \nu) \left[\sum_{i=1}^{2} w_{ik} \left(1 + \sum_{s=1}^{8} \gamma^k_{is} \left(x^s_{it} - \mu_{is} \right)^k \right) \right],\ k = 1, 2$

	MGC_1	MGC_2	$MGME^+_1$	$MGME^+_2$	$MGME^*_1$	MST
			Panel 1: Portfolio FX \$/£ - FX €/£			
Stage 1						
ϕ_{10}	-.002 (-0.25)*				-.002 (-0.26)*	-.002 (-0.24)*
ϕ_{11}	.070 (4.20)				.070 (4.19)	.007 (4.21)
α_{10}	.001 (2.14)				.001 (1.33)*	.001 (1.28)*
α_{11}	.034 (5.89)				.029 (5.13)	.009 (4.41)
α_{12}	.962 (87.49)				.973 (84.34)	.983 (81.73)
ϕ_{20}	-.001 (-0.24)*				-.002 (-0.25)*	-.002 (-0.23)*
ϕ_{21}	.033 (2.02)				.034 (2.00)	.033 (2.10)
α_{20}	.001 (1.67)*				.001 (1.46)*	.001 (1.41)*
α_{21}	.023 (6.16)				.033 (5.30)	.024 (4.89)
α_{22}	.971 (99.08)				.962 (85.84)	.974 (86.80)
Stage 2						
γ_{12}	-.017 (-0.62)*	.091 (2.91)	-.463 (3.83)	-.121 (-5.28)		.099 (1.40)*
γ_{14}	.073 (4.60)		.141 (3.38)			.149 (7.17)
γ_{16}	.009 (1.92)*		-.017 (-1.84)*			
γ_{18}	.001 (3.39)	.001 (3.81)	.008 (1.93)*	-.001 (-1.65)*		
γ_{22}	-.024 (-0.87)*	.096 (3.46)	-.776 (-6.18)	.105 (4.25)		.070 (0.91)*
γ_{24}	.065 (4.19)		.301 (4.86)			.003 (2.65)
γ_{26}	.007 (1.64)*		-.039 (-3.78)			
γ_{28}	.001 (1.93)*	-.001 (-2.91)	.001 (3.40)	.001 (2.10)		
ν					8.10 (11.20)	13.1 (11.29)
ρ	.273 (17.72)	.298 (18.57)	.275 (17.72)	.292 (18.13)	.271 (15.74)	.172 (12.94)
LnL	-2292.60	-2330.14	-2292.60	-2312.56	-4240.58	-4228.87
BIC	2329.35	2350.56	2329.35	2332.97	4273.25	4277.87

TABLE 2 (continued)

	MGC_1	MGC_2	$MGME_1^+$	$MGME_2^+$	$MGME_1^*$	MST
			Panel 2: Portfolio FX Y/£ - FX y/£			
Stage 1						
ϕ_{10}		-.008 (-0.90)*			-.001 (-0.67)*	-.008 (-0.92)*
ϕ_{11}		.009 (0.54)*			.009 (0.55)*	.009 (0.54)*
α_{10}		.002 (1.67)*			.000 (0.03)*	.001 (0.29)*
α_{11}		.028 (5.33)			.012 (11.30)	.005 (3.70)
α_{12}		.977 (77.42)			.987 (84.43)	.994 (75.16)
ϕ_{20}		-.000 (-0.76)*			-.001 (-0.77)*	-.001 (-0.76)*
ϕ_{21}		.044 (2.63)			.044 (2.75)	.045 (2.06)
α_{20}		.007 (4.02)			.010 (3.66)*	.004 (3.31)
α_{21}		.066 (9.02)			.077 (7.00)	.033 (5.49)
α_{22}		.930 (93.24)			.911 (62.84)	.908 (60.66)
Stage 2						
γ_{12}	-.022 (-0.78)*	.116 (4.90)	-1.08 (8.85)	-.133 (-6.47)		-.061 (-0.94)*
γ_{14}	.108 (5.96)		.435 (6.83)			.191 (10.87)
γ_{16}	.016 (3.14)		-.059 (-5.70)			
γ_{18}	.002 (5.15)	-.001 (-4.04)	.002 (5.16)	-.001 (-2.91)		
γ_{22}	-.035 (-1.06)*	.115 (4.71)	-1.06 (-8.67)	.156 (8.75)		.163 (2.94)
γ_{24}	.162 (7.27)		.362 (5.63)			.047 (4.28)
γ_{26}	.023 (3.92)		-.049 (-4.72)			
γ_{28}	.002 (4.76)	-.001 (-6.12)	.002 (4.76)	-.001 (3.03)		
ν					5.89 (14.39)	10.2 (17.01)
ρ	.041 (2.55)	.054 (2.76)	.041 (2.55)	.058 (3.08)	.040 (2.13)	.024 (2.29)
LnL	-4142.84	-4268.56	-4142.84	-4202.82	-6078.37	-6041.16
BIC	4179.59	4288.98	4179.59	4223.24	6111.03	6090.16

Notes: This table presents (Q)ML estimates of the parameters of the MGME densities using MN ($MGME_k^+$) and MST ($MGME_k^*$) as basis, $k = 2$ indicates that positivity transformations are implemented. The estimates of two versions of the MGC and the MST are also displayed, all distributions for the bivariate case. ϕ_{is} $(s = 0, 1)$ stand for the AR(1) parameters of the conditional means and α_{is} $(s = 0, 1, 2)$ for the GARCH(1,1) parameters of the conditional variances. ρ denotes the unconditional correlation parameter, γ_{is} $(s = 2, 4, 6, 8)$ the $s - th$ order polynomial weighting parameter of the expansions and ν the degrees of freedom of the MST. The log likelihood value (LnL) and the Schwarz Bayesian Information Criterion (BIC) are displayed in the last two rows.

Figure 1. Bivariate GME density (yuan/pound-yen/pound)

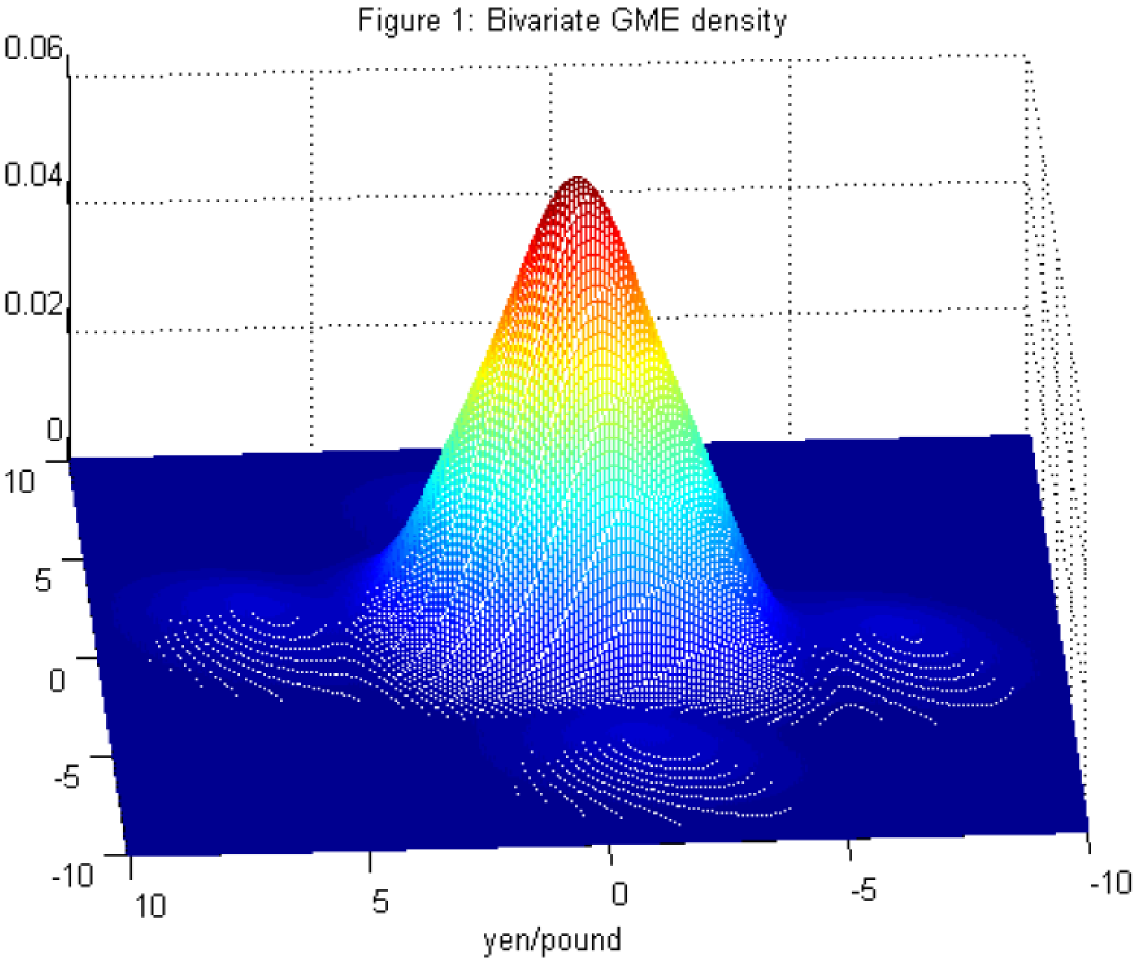

Figure 2A: Marginal densities (yen/pound)

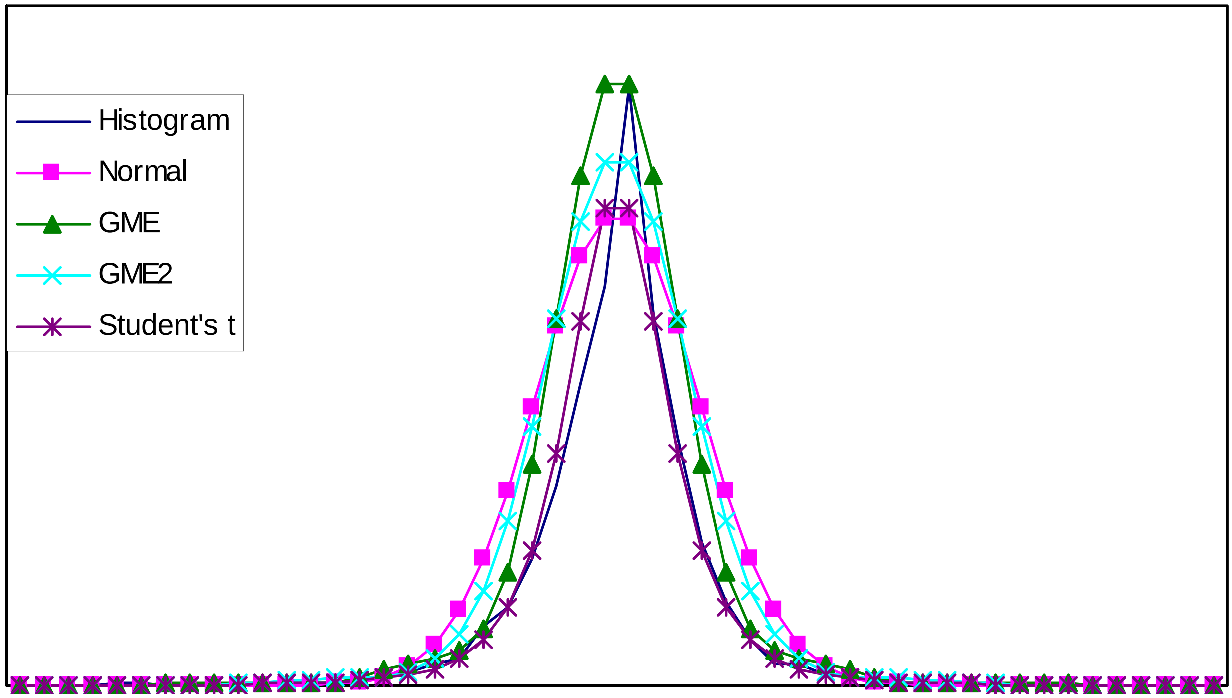

Figure 2B: Left talils of marginal densities (yen/pound)