Multivariate Distributions based on Moments Expansions and Hermite Polynomials Approximations

Trino-Manuel Ñíguez

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

Javier Perote*

Department of Economics, University of Salamanca, Salamanca, Spain

Abstract

This paper proposes a new semi-nonparametric (SNP) method for the modeling of the multivariate distribution of portfolio returns. The distribution we obtain, named Multivariate Moments Expansion (MME), is specified using expansions of a parametric density through a novel type of polynomials. A statistical analysis of the MME shows that it is as flexible as other SNP methods being simpler to implement and more tractable theoretically. We demonstrate that Gaussian MME formally admits DCC-type two-stage maximum-likelihood estimation. An in-sample application for foreign exchange porfolios provides evidence on MME improvements in goodness-of-fit compared to other multivariate SNP methods.

Keywords: Dynamic Conditional Correlation; Gram-Charlier series; Multivariate GARCH; Semi-nonparametric methods.

JEL classification: C16, G1.

*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 portfolio risk of has traditionally been tackled through multivariate volatility and GARCH (MGARCH) models, literature reviewed comprehensively in Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Terasvirta (2009). The recent research on these models has focused on the problem of dimensionality together with: (i) better explaining time-varying correlations; see Engle (2002), Cappiello, Engle and Sheppard (2007), and Engle and Kelly (2012) for parametric specifications; and Harris and Mazibas (2013), and Aslandis and Casas (2013) for semi-parametric and non-parametric specifications, respectively; (ii) better understanding the properties of the estimation; see Sentana, Calzolari and Fiorentini (2008) and (iii) better accounting for the empirical features of the distribution of returns.

Regarding the latter, the Normal distribution is widely assumed, in spite of the non-normality of returns, due to its well-known statistical properties. Namely, in standard correctly specified MGARCH models with normal errors, consistent (quasi)-maximum likelihood estimators can be obtained, while the estimation properties under non-normality are not known. Consistency is thus achieved under normality at the expense of losses in goodness-of-fit and forecasting performance for higher-order moments, and even for volatility in particular for periods of high uncertainty (see Del Brio, Ñíguez and Perote, 2011). The alternative distributions to the multivariate Normal used in the literature to address the later point include the following:

1. Parametric multivariate distributions such as the Skewed Normal (Azzalini and Dalla Valle, 1996), Student's t (Kotz and Nadarajah, 2004), Weibull (Malevergne and Sornette, 2004), Kotz-type (Olcay, 2005) or the Generalized Hyperbolic (Fajardo and Farias, 2010).

2. Copulas as a straightforward method to construct an implicit multivariate probability density function (pdf) from any combination of univariate marginals; see Patton, 2012, for a survey of copula models in econometrics. This approach, however, yields pdfs that are analytically less tractable for financial applications that require computationally demanding numerical algorithms, such as Value-at-Risk (VaR) forecasting.

3. Semi-nonparametric (SNP) densities such as the multivariate Gram-Charlier pdfs in Del Brio, Ñíguez and Perote (2009) have been shown to accurately account for fat tails and skewness. More recently, in Del Brio et al. (2011) a generalization

to the SNP framework of the Dynamic Conditional Correlation (DCC) method of Engle (2002) is proposed resulting in the first non-Normal DCC model (SNP DCC), which formally admits two-stage maximum-likelihood (ML) estimation. This SNP DCC model can capture skewness and kurtosis as well as the dynamics of large variance-covariance matrices. SNP pdfs present also well-known issues related to the truncation needed in empirical applications, i.e., a well-defined truncated SNP pdf requires conditions to ensure positiveness in its whole parametric space. This problem is usually addressed through Gallant & Nychka (1987)- and Gallant & Tauchen (1989)-type of transformations (GNT hereafter).[1] However, this solution introduces non-linear relations among the distribution moments that result in complex (positive) GNT to handle theoretically.

In this paper we propose a new multivariate SNP family of densities, called multivariate Moments Expansion (MME), that provides a solution to the latter issue. We show that MME present a simpler specification compared to other SNP densities that allows to address the complexity inherent to GNT SNP pdfs, preserving the good properties of Gram Charlier (MGC) pdfs. MME pdfs exhibit the following econometric features:

1. *Flexibility*: they preserve the flexibility proper of GC series to adapt to the shape of any target pdf.

2. *Simplicity:* they only require that the distribution used as basis has finite moments up to the truncation order.

3. *Generality*: non-GNT-transformed MGC pdfs can be obtained as particular cases.

4. *Positiveness*: they straightforwardly admit GNT-type of re-specifications to guarantee positiveness.

5. *Tractability*: when the Gaussian pdf is used as basis they formally admit the DCC-type of separability (variances-correlations) of the likelihood function.

6. *Applicability*: they present a simple formulation that facilitates the model implementation, e.g. admitting straightforwardly the specification of (un)conditional (higher-order) (co)-moments with possible asymmetries.

[1] See León, Mencía and Sentana (2009), and Ñíguez and Perote (2012) for applications of this solution to univariate Gram-Charlier distributions. An alternative solution consists on constraining the parametric space (Jondeau and Rockinger, 2001) at the cost of limiting the goodness-of-fit of the distribution.

The remaining of the article is organized as follows. In Section 2, we define the MME pdf and discuss its statistical properties emphasizing on the econometric features above. Section 3 provides an empirical application for bivariate portfolios of exchange rates (FX hereafter) to test the applicability and goodness-of-fit of our proposed pdf. Section 4 summarizes our conclusions.

2 The Multivariate Moments Expansion Distribution

In this section we present the MME pdf firstly defined for the case of uncorrelated variables and secondly for the more general case of correlated variables. To provide a clear overview of the advantages of the MME density, we will analyze it compared to those of the MGC in relation to: (a) the polynomial structure, (b) the distribution being expanded and (c) the GNT reformulations effects on the specification.

2.1 MME for uncorrelated variables

The generalization of a univariate SNP density to the n-dimensional setting may be done through two stages. In the first stage, the multivariate pdf is constructed as the product of n independent SNP distributed variables. In the second stage, a non-diagonal correlation matrix is incorporated by means of a linear transformation.[2] However, MGC pdfs obtained by this two-stage procedure present problems related to both parameter identification and the estimation convergence due to the pdf number of parameters, which increases dramatically with the expansion (m) and dimension (n) orders.

In this paper we use an approach to generalize SNP pdfs to the multivariate framework that differs to the usual two-stage procedure in the stage 1. Thus, stage 1 involves the characterization of the multivariate SNP of uncorrelated variables through a multivariate pdf with number of parameters increasing linearly with the dimension n, and whose marginals are distributed as univariate SNP distributions. It is noteworthy that the family of SNP distributions constructed through this procedure are then uncorrelated but not independent (with the exception of the multivariate Gaussian distribution).

[2] Alternatively, the product of marginals might be multiplied by SNP copulas obtaining different multivarite SNP distributions, which may also account for tail dependence. We leave this approach for further research.

For the sake of using a consistent and general notation, our definitions include both the standard (non-positive) case of SNP pdfs, denoted by $k=1$, and its (positive) GNT pdfs counterparts, denoted by $k=2$.[3]

In addition, we consider two families of SNP distributions: The first one (Definition 1), denoted as Π_k, is defined on the Gaussian pdf derivatives and the so-called Hermite polynomials (HP) and gathers the MGC as special case. The second family that we name MME, (Definition 2), denoted as F_k, is defined for a general sequence of distributions used as basis and involve very simple polynomial structures that depend directly on the non-central moments of the sequence of basis distributions. Without loss of generality, we consider the same order for the expansion of every dimension, m, in all cases.

The remaining of this section is devoted to formally define these two pdf families and to analyze their relation and properties for the uncorrelated case. In the next section, we study the case of their correlated pdf counterparts.

Definition 1. *Let $\{x_{it}\}_{i=1}^{n}$ be a sequence of uncorrelated standard normal variables. Then, we define the joint pdf Π_k as,*

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

where $\mathbf{x}_t=(x_{1t},x_{2t},\ldots,x_{nt})'\in\mathbb{R}^n$, $\phi\left(x_{it}\right)$ denotes the standard Gaussian pdf, $\boldsymbol{\Delta}$ is a matrix of parameters with general element $\{\delta_{si}\}$ and $\{\lambda_{ik}\}_{i=1}^{n}$ is the sequence of scaling constants that makes the density to integrate up to one

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

The SNP pdf Π_k nests the MGC when $k=1$ and the GNT MGC when $k=2$. Given the constants in equation (2), both Π_1 and Π_2 integrate one, but only Π_2 is a well-defined pdf (i.e. positive) in the whole parameter domain. The statistical properties of both densities can be found in Del Brio et al. (2011); see León et al. (2009) for other related GNT GC

[3] For the sake of simplicity we only assume the simplest GNT transformation, i.e. squaring every single polynomial in the SNP expansion (see Ñíguez and Perote, 2012). This transformation yields symmetric pdfs, unless the distribution used as basis is skewed.

specifications. It is noteworthy that the Π_k marginals are distributed as univariate GC and their moment structure may be easily worked out since the HPs form an orthonormal basis. Particularly, for every marginal i in Π_1 the non-central moment of order r depend on $\{\delta_{si}\}$ with $s \leq r$. Furthermore, for Π_1 the even/odd moments depend only on the even/odd parameters.[4] However, the moment structure for the GNT GC (Π_2) is much more complex, i.e., the moments of all order depend on all density parameters and there are nonlinearity between the density parameters and moments – see Ñíguez and Perote (2012) for further details on this point.

Next we define the MME pdfs and compare them with the Π_k pdfs (Definition 1), highlighting the contribution of the former.

Definition 2. *Let $\{\varphi_i(x_{it})\}_{i=1}^n$ be a sequence of pdfs of uncorrelated variables with $E[x_{it}^r] = \int x_{it}^r g_i(x_{it})dx_{it} = \mu_{ir} < \infty$ $\forall i = 1, 2, \ldots, n$ and $\forall r \leq m$, m being the expansion order. Then, the MME pdf of $\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{nt})' \in \mathbb{R}^n$ is defined as,*

$$F_k(\mathbf{x}_t; \boldsymbol{\gamma}) = \frac{1}{n}\left[\prod_{i=1}^{n}\varphi_i(x_{it})\right]\left[\sum_{i=1}^{n} w_{ik}^{-1}\left(1 + \sum_{s=1}^{m}\gamma_{is}^k (x_{it}^s - \mu_{is})^k\right)\right], \ \forall k = 1, 2 \tag{3}$$

where $\boldsymbol{\gamma}$ is a matrix of parameters with general element $\{\gamma_{is}\}$, and $\{w_{ik}\}_{i=1}^n$ is the sequence of scaling constants that makes the density to integrate up to one

$$w_{ik} = \int \left[1 + \sum_{s=1}^{m}\gamma_{is}^k (x_{it}^s - \mu_{is})^k\right] g_i(x_{it})dx_{it} = \begin{cases} 1 & \text{if } k = 1 \\ 1 + \sum_{s=1}^{m}\gamma_{is}^2(\mu_{i,2s} - \mu_{is}^2) & \text{if } k = 2 \end{cases} \quad \forall i = 1, 2, ..., n. \tag{4}$$

The pdf F_2 is well-defined – it integrates one (see *Proof* 1 in the Appendix 1) and is positive for the constants w_{i2} in equation (4) – but F_1 needs further restrictions to ensure positivity (see Níguez and Perote, 2013). We leave the discussion on the properties of the MME for the next section, where marginals (*Proof* 2), non-central moments (*Proof* 3) and cdfs (*Proof* 4) are derived, to focus on the comparison between the SNP densities in Definition 1 and 2. The differences between Π_k and F_k come basically from two main sources: The distributions used as basis and the polynomial structure in the expansions. On the one

[4]For instance, δ_{1i} and δ_{2i} capture the unconditional mean and variance of x_{it}, respectively, and δ_{3i} and δ_{4i} the unconditional skewness and excesss kurtosis, respectively, provided that $\delta_{1i} = \delta_{2i} = 0$

hand, F_k is defined for any sequence $\{\varphi_i(x_{it})\}_{i=1}^n$ of distributions with the only assumption of having finite moments up to the expansion order $m, \forall i = 1, 2, ..., n$ and its polynomials have a very simple form: $P_s(x_{it}) = x_{it}^s - \mu_{is}$. On the other hand, Π_k is only defined for expansions of the Normal pdf, $\phi(x_{it})$, and then the expansions are in terms of the HPs, $H_s(x_{it})$, which structure depend on the derivatives of the Normal pdf. The HPs form an orthonormal basis and this fact is the key property to interpret the GC truncated series as a pdf. The polynomials of the MME, however, are not orthogononal but their simple structure is enough to guarantee up to one integration. Furthermore, when $\{\varphi_i(x_{it})\}_{i=1}^n$ is a the sequence of standard Gaussian distributions with $s-th$ order non-central moments denoted by μ_{is}^+, then $H_s(x_{it})$ can be rewritten as a linear combination of $\left\{x_{it}^s - \mu_{is}^+\right\}$ as follows

$$H_s(x_{it}) = s! \sum_{k=0}^{[s/2]} \frac{(-1)^k}{k!(s-2k)!2^k} x_{it}^{s-2k} = s! \sum_{k=0}^{[s/2-1]} \frac{(-1)^k}{k!(s-2k)!2^k} \left[x_{it}^{s-2k} - \mu_{s-2k}^+\right]. \tag{5}$$

Therefore if the standard Gaussian is used as basis and $k = 1$, then both F_k and Π_k pdfs are reparametrisations of the MGC, i.e. $F_1^+(\mathbf{x}_t; \boldsymbol{\gamma}) = \Pi_1(\mathbf{x}_t; \boldsymbol{\Delta})$. Nevertheless when the standard Gaussian is used as basis but $k = 2$, then F_k and Π_k do not render the same pdf, i.e. $F_2^+(\mathbf{x}_t; \boldsymbol{\gamma}) \neq H_2(\mathbf{x}_t; \boldsymbol{\Delta})$. Therefore G_k and Π_k are not nested but they both nest the MGC (provided that $\varphi_i(x_{it}) = \phi(x_{it})$ $\forall i = 1, 2, ..., n$ and $k = 1$). This means that the pdfs in Definitions 1 and 2 are both valid asymptotic expansions, but also that when GNT transformations are implemented the MME represents a much simpler specification.

Another important difference between F_k and Π_k pdfs lies in the fact that the former naturaly admits expansions not only for the Gaussian density but also for any other distribution with enough (at least up to the expansion order m) finite non-central moments.[5] This fact opens a hopefully fruitful line of research on SNP methods. For example, in the Appendix 2 we provide the density of a MME expansion of the multivariate Student's (MST) with v degrees of freedom (df), referred as F_k^*. This pdf is well defined provided that $v > m$ and converges to the MGC as v tends to infinity. Therefore F_k^* represents a straightforward generalization of the MGC.

These arguments reveal the MME as a very simple, general and tractable SNP method to parsimoniously approximating the portfolio return distribution. The advantages of the

[5]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 these SNP pdfs are the so-called Gram-Charlier Type B, Laguerre and Jacobi expansions, respectively. Nevertheless their empirical applications and their extensions to the multivariate framework are unusual, likely for tractability reasons.

F_k over Π_k are particularly useful when GNT transformations are implemented ($k = 2$), since in this case F_2 is much more tractable than Π_2 from both the theoretical and empirical viewpoint.

Next we proof some interesting properties of the MME pdf and in Section 3 we compare the performance of the MME compared to other SNP alternatives.

2.1.1 MME properties

This Section states three main properties of the MME: the marginals, the non-central moments and the cdfs of both F_1 and F_2. The proofs of the properties are provided in the Appendix 1 for the sake of simplicity in the exposition.

1. Marginal densities (Proof 2 in the Appendix 1).

$$f_{ki}(x_{it}) = \begin{cases} g_i(x_{it}) \left[1 + \sum\limits_{s=1}^{m} \dfrac{\gamma_{is}}{n} \left(x_{it}^s - \mu_{is} \right) \right], & \text{if } k = 1, \\ g_i(x_{it}) \left[\dfrac{n-1}{n} + \dfrac{w_{i2}}{n} \left(1 + \sum\limits_{s=1}^{m} \gamma_{is}^2 \left(x_{it}^s - \mu_{is} \right)^2 \right) \right], & \text{if } k = 2. \end{cases} \tag{6}$$

2. Non-central moments (Proof 3 in the Appendix 1).

$$E_k \left[x_{it}^r \right] = \begin{cases} \mu_{ir} + \sum\limits_{s=1}^{m} \frac{\gamma_{is}}{n} (\mu_{i,r+s} - \mu_{ir}\mu_{is}), & \text{if } k = 1, \\ \left[\frac{n-1}{n} + \frac{w_{ik}}{n} \right] \mu_{ir} + \frac{w_{ik}}{n} \sum\limits_{s=1}^{m} \gamma_{is}^2 \left[\mu_{i,2s+r} + \mu_{is}(\mu_{is}\mu_{ir} - 2\mu_{i,s+r}) \right], & \text{if } k = 2, \quad \forall r \in \mathbb{N}. \end{cases} \tag{7}$$

3. Cumulative distribution function (Proof 4 in the Appendix 1).

$$\Pr[x_1 \leq \overline{x}_1, \cdots, x_n \leq \overline{x}_n]_k = \frac{1}{n} \sum_{i=1}^{n} h_{ik}(\overline{x}_i) \left[\prod_{j=1,\, j \neq i}^{n} \int_{-\infty}^{\overline{x}_j} g_j\left(x_{jt}\right) dx_{jt} \right], \; \forall k = 1, 2. \tag{8}$$

where $h_{ik}(\overline{x}_i)$ stands for the cdf of the corresponding univariate ME distribution evaluated at $\overline{x}_i$ (see Ñíguez and Perote., 2014, for a closed formula of $h_{ik}(\overline{x}_i)$ for the Gaussian case).

2.2 MME for Correlated Variables

So far we have analysed the case of the MME distribution of uncorrelated variables (Definitions 1 and 2). However, dependencies among the variables are incorporated in the MME density in Definition 3 by considering a linear transformation of the type

$$\mathbf{u}_t = \mathbf{\Sigma}_t^{1/2} \mathbf{x}_t = \mathbf{D}_t \mathbf{R}^{1/2} \mathbf{x}_t, \tag{9}$$

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, $\mathbf{R}$.[6]

Definition 3. *Let $\mathbf{u}_t = (u_{1t}, u_{2t}, \ldots, u_{nt})' \in \mathbb{R}^n$ a 0 mean random vector with multivariate pdf $G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta})$, where $\mathbf{\Sigma}_t$ stands for the conditional variance and covariance matrix and $\boldsymbol{\theta}$ being a vector containing the rest of the density parameters. Let $g_i(u_{it})$ denote the $i-th$ marginal density of $G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta})$ and $\mu_{ir} < \infty$ $\forall i = 1, 2, \ldots, n$ and $\forall r \leq m$ its corresponding $s-th$ order non-central moment. The MME pdf of $\mathbf{u}_t$ is defined as,*

$$F_k(\mathbf{u}_t; \boldsymbol{\gamma}, \mathbf{\Sigma}_t, \boldsymbol{\theta}) = \frac{1}{n} G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta}) \sum_{i=1}^{n} w_{ik}^{-1} \left[1 + \sum_{s=1}^{m} \gamma_{is}^k \left(x_{it}^s - \mu_{is} \right)^k \right], \ \forall k = 1, 2, \qquad (10)$$

where $\boldsymbol{\gamma}$ is a matrix of parameters with general element $\{\gamma_{is}\}$, x_{it} is the corresponding component of the inverse transformation in equation (9) and w_{ik} is the scaling constant in equation (4).

The MME in equation (10) nests not only the multivariate distribution used as basis, $G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta})$, but also expansions of whichever pdf nested in $G(\mathbf{u}_t; \mathbf{\Sigma}_t, \boldsymbol{\theta})$. An interesting case is the MME expansion of elliptical distributions, which is detailed in the Appendix 2. This SNP family encompass both the MGC (F_k^+) and the expansion of the MST (F_k^*), as well as GNT transformations of these cases for $k = 2$.

The statistical properties of the MME family of distributions in Definition 3 may be straightforwardly obtained from those of the uncorrelated case in equations (3) and (4), and taking into account the transformation in (9) and the properties of the pdfs used as basis. Interestingly, the MME distributions that use Gaussian pdfs 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 (11), and the log-likelihood of the MME expressed 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 (12), where $\boldsymbol{\alpha}$, $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ stand for the matrices containing the parameters of the conditional variances, correlations and expansion

[6] Without loss of generality we are considering constant correlations among the variables, but the DCC model can be also directly implemented.

terms of the MME distribution, respectively (see Proof 5 in Appendix 1).

$$
\begin{aligned}
L_V(\mathbf{u}_t, \boldsymbol{\alpha}) &= -\frac{1}{2}\sum_{i=1}^{n}\left[T\log(2\pi) + \sum_{t=1}^{T}\left(\ln(\sigma_{it}^2) + \frac{u_{it}^2}{\sigma_{it}^2}\right)\right] \\
&= -\frac{1}{2}\sum_{i=1}^{n}\left[T\log(2\pi) + L_{V_i}(\boldsymbol{\alpha}_i)\right], \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
L_{F_k^+}(\boldsymbol{\varepsilon}_t, \boldsymbol{\alpha}, \boldsymbol{\rho}, \boldsymbol{\gamma}) &= -\tfrac{1}{2}\sum_{t=1}^{T}\left\{\ln|\mathbf{R}| + \boldsymbol{\varepsilon}_t'\mathbf{R}^{-1}\boldsymbol{\varepsilon}_t - 2\ln\left[\sum_{i=1}^{n} w_{ik}\left(1 + \sum_{s=1}^{m}\gamma_{is}^k\left(x_{it}^s - \mu_{is}^+\right)^k\right)\right]\right\} \\
\forall k &= 1, 2. \qquad (12)
\end{aligned}
$$

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 jointly estimated with the rest of the MME parameters by limited information ML (LIML) applied to the log-likelihood concentrated on the estimates of the conditional variances, $L_{F_k^+}(\boldsymbol{\varepsilon}_t, \widehat{\boldsymbol{\alpha}}, \boldsymbol{\rho}, \boldsymbol{\gamma})$ where $\widehat{\boldsymbol{\alpha}} = \arg\max\{L_V(\mathbf{u}_t, \boldsymbol{\alpha})\}$ (second step).[7] Such a procedure simplifies by far the estimation compared to the jointly estimation, which is the only theoretically valid ML methodology if the assumed distribution is not Normal. Del Brio et al. (2011) extended this two-step methodology to non-Gaussian distributions of the type displayed in equation (1). These authors argued that as the GC is a valid asymptotic expansion the second step is consistent even under mispecification. The same argument holds for our MME approach, since it encompasses the MGC distribution.

3 Application to exchange rate portfolios

3.1 The model

Let us consider a portfolio of n assets and let $\mathbf{r}_t$ the vector of the returns on these assets at time t. We assume that the distribution of $\mathbf{r}_t$ conditioned on the information set $\boldsymbol{\Omega}_{t-1}$ belongs to the MME family with conditional mean, $\mathbf{E}(\mathbf{r}_t|\boldsymbol{\Omega}_{t-1}) = \boldsymbol{\mu}_t(\boldsymbol{\phi})$, and and conditional variance and covariance matrix $\mathbf{E}(\mathbf{u}_t\mathbf{u}_t'|\boldsymbol{\Omega}_{t-1}) = \boldsymbol{\Sigma}_t(\boldsymbol{\alpha}, \boldsymbol{\rho})$ (equations (13), (14) and (15)), where $\boldsymbol{\phi}$, $\boldsymbol{\alpha}$, $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ are the matrices including conditional mean, variance, correlation

[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.

and MME expansion parameters, respectively. We model conditional mean and variance for every variable as AR(1)-GARCH(1,1) (equations (16) and (17), $\circ$ 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]

$$\begin{aligned}
\mathbf{r}_t &= \boldsymbol{\mu}_t(\boldsymbol{\phi}) + \mathbf{u}_t, && (13)\\
\mathbf{u}_t|\Omega_{t-1} &\sim F_k(\mathbf{0}, \boldsymbol{\Sigma}_t(\boldsymbol{\alpha}, \boldsymbol{\rho}), \boldsymbol{\gamma}), \text{ or } \mathbf{u}_t|\Omega_{t-1} \sim H_k(\mathbf{0}, \boldsymbol{\Sigma}_t(\boldsymbol{\alpha}, \boldsymbol{\rho}), \boldsymbol{\gamma}), && (14)\\
\boldsymbol{\Sigma}_t(\boldsymbol{\alpha}, \boldsymbol{\rho}) &= \mathbf{D}_t(\boldsymbol{\alpha})\mathbf{R}(\boldsymbol{\rho})\mathbf{D}_t(\boldsymbol{\alpha}), && (15)\\
\boldsymbol{\mu}_t(\boldsymbol{\phi}) &= \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1'\mathbf{r}_{t-1} && (16)\\
\mathbf{D}_t^2 &= diag\{\alpha_{i0}\} + diag\{\alpha_{i1}\} \circ \mathbf{u}_{t-1}\mathbf{u}_{t-1}' + diag\{\alpha_{i2}\} \circ \mathbf{D}_{t-1}^2, && (17)
\end{aligned}$$

We consider different SNP models that are nested in the MME family of distributions. In particular we consider four different expansions of the Gaussian pdf: Π_k (equation (1)) and F_k^+ (equation (21)) $\forall k = 1, 2$. We choose these specifications to show that the MGC (Π_1) is nested in the the MME family (F_1^+) and that its GNT transformation (F_2^+) may be a more simple and accurate formulation than the Π_2. Furthermore, in order to show the potential advantages of the MME when using other densities as basis we also include a version of the MME espansion of the Student's t, F_1^* (equation (22)), which we compare with a version of the multivariate Student's t (MST) as benchmark.[9] In the Appendix 2 we detail all these cases belonging to the MME family that uses elliptical distributions as basis. 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 MME densities using Gaussian pdfs as basis to larger portfolios, however, are not computationally demanding since they can be consistently estimated in two (or three) steps.

For the bivariate case and using the symmetric eigenvalue decomposition of $\boldsymbol{\Sigma}_t$, the inverse transformation of equation (9) becomes

[8] We choose the more benchmarked models for first and second moments since we focus on the advantages of the MME distribution compared to other multivariate SNP alternatives. Obviously, our model admits more complex multivarite volatilitiy structures such as the DCC.

[9] There exist many other specifications for the multivariate Student's t (see Kotz & Nadarajah, 2004).

$$x_{it} = \frac{1}{2}\sum_{j=1}^{2}\left(\frac{1}{\sqrt{1+\rho}} + (-1)^{i+j}\frac{1}{\sqrt{1-\rho}}\right)\frac{u_{jt}}{\sigma_{jt}},\ \forall i = 1,2, \tag{18}$$

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}.[10]

3.2 The Data

In this section we investigate the empirical performance of the six bivariate models discussed above: Π_1, Π_2, F_1^+, F_2^+, F_1^* and MST. The data used are (daily) percent log returns computed as $r_{it} = 100\log(P_{it}/P_{it-1})$ from series $\{P_{it}\}_{t=1}^{T}$ 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 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 MME models that expand the multivariate Normal (MN) is carried out in two stages by (Q)ML 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

[10] 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. Furthermore, 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 included in the multivariante distribution used as basis, $G(\mathbf{u}_t;\mathbf{\Sigma}_t,\boldsymbol{\theta})$.

to estimate correlations and the parameters of the expansion. Robust standard errors were computed following Bollerslev and Wooldridge (1992).[11] The models that include the MST are jointly estimated by ML since two step procedures cannot be theoretically applied because the log-likelihood is not separable.[12]

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 MME 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 weight parameter of the expansions and ν for the df of the MST (note that the estimates of the expansions are well defined since $v > 8$). T-ratios for robust standard errors are in parentheses next to the parameter estimates and an asterisk signals the insignificant parameters at 5 percent confidence level.

Regarding the specification of the (GNT) MGC and (GNT) MME 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 (H_2 and F_2^+) and parameters γ_{i6} and γ_{i8} $(i = 1, 2)$ of the expansion of the MST (F_1^*) are also omitted. The selection of the parametric structure for these expansions was done according to linear restriction (Wald) tests but there exists a clear intuition behind these particular parametric structures, since the parameters of the MME 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 thick tails and thus a larger expansion is not required.

[11] See Del Brio et al. (2011) for a discussion on the properties and implementation of 2-stage estimation methods for SNP models.

[12] Nevertheless when multi-step procedures are implemented to the MST, the estimation results do not differ significantly from the one-step approach (Bauwens and Laurent, 2005, and Jondeau and Rockinger, 2005).

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.[13] According to these criteria we observe the following evidence: (1) The MME models provide a notably better goodness-of-fit than the distributions used as basis by themselves (Gaussian[14] or Student's t). (2) If GNT transformations are not implemented the MGC and MME yield the same density, although the latter specification seems to increase parameters significance (i.e. 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 GNT 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 GNT MME seems to provide a better fit than the traditional GNT SNP, although both specifications are not nested. (5) The expansion of the MN requires additional terms to capture fat tails compared to the MST, since the latter 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) does not necesarily yield consistent estimates if two-step procedures are implemented.

We also observe the usual small structure in the conditional mean, high persistence in the conditional variance and the correlation parameters of \$/£ and the €/£ FX are higher 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 ME 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

[13] 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 and Meese, 1981).

[14] 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.

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 ME densities clearly outperform the Normal and the Student's t (specially at the tails) and the ME (ME) presents a more accurate data fit than the GNT ME (ME2).

[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. Parametric distributions or copulas has been proposed for that purpose, in spite of the fact that they 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. Alternatively, the SNP approach based on MGC expansions allows to straightforwardly address these issues since, not only they can fit any target density through their general and flexible parametric structure, but also because they present an analytical specification that is tractable due to the orthogonal structure of Hermite polynomials. In this paper we propose a novel SNP family of distributions, the MME, that are much simpler (the polynomials of the expansion do not require orthogonality) and thus easier to implement, even when used to expand any distribution as it only requires having as many finite moments as the expansion length, for every dimension. Furthermore, our approach generalizes the univariate ME proposed by Ñíguez and Perote (2014) to a multivariate framework.

If the Gaussian density is used as basis and positive transformations are not implemented, the MME is just a re-specification of the MGC. But if positivity is imposed, the MME yields a simpler and more accurate formulation. In these cases, the MME also admits the decomposition of the likelihood function proposed in Engle (2002), which permits estimating independently the volatility processes of every asset (stage 1 ML) and, in a second stage, the rest of the density parameters (LIML) (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 MME pdfs for modeling the distribution of FX portfolio returns ($/£-€/£ and Y/£-y/£ FX), in relation to the MN and MST, taken as benchmark, and alternative MGC pdfs. The results show that the MME specifications outperform not only the MN and MST but also might be superior to the MGC pdfs when positive transformations are implemented, thus the MME poses as an interesting and easy-to-implement SNP tool for modeling portfolio distributions.

Appendix 1

This appendix includes the proofs of the properties of the MME densities presented in Section 2.1. *Proof 1* shows that MME 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 MME.

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.

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

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.

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

Proof 5.

$$
\begin{aligned}
L_{F_k^+}(\mathbf{u}_t,\boldsymbol{\alpha},\boldsymbol{\rho},\boldsymbol{\gamma}) &= -\frac{1}{2}\sum_{t=1}^{T}\left(n\ln(2\pi)+\ln|\boldsymbol{\Sigma}_t|+\mathbf{u}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{u}_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]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(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\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]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(n\ln(2\pi)+2\ln|\mathbf{D}_t|+\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]-T\ln(n) \\
&= -\frac{1}{2}\sum_{t=1}^{T}\left(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\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]-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}
$$

Appendix 2: The case of the MME for elliptical pdfs

In this Appendix we provide the MME for the special case in which elliptical pdfs are used as basis and for a variance and covarianve matrix $\boldsymbol{\Sigma}_t$. The distributions considered in the empirical application in Section 3 are nested in this particular family of SNP distributions.

Let us consider the following eliptical family of distributions

$$G(\mathbf{u}_t;\boldsymbol{\Sigma}_t,\boldsymbol{\theta})=|\boldsymbol{\Sigma}_t|^{-1/2}\,\varphi_n\left(\mathbf{u}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{u}_t\right), \tag{19}$$

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

$$\int_0^{\infty} z^{n/2-1} \varphi_n(z)\, dz = \frac{\Gamma(\frac{n}{2})}{\pi^{n/2}}. \tag{20}$$

Therefore if $\varphi_n(z) = \frac{e^{-z/2}}{(2\pi)^{n/2}}$ the MME that expands the Gaussian pdfs is given in equation (21) and if $\varphi_n(z) = \left(1 + \frac{z}{\nu}\right)^{-(\nu+n)/2}$, ν being the degrees of freedom, the expansion of the Student t pdfs is that of equation (22)).

$$\begin{aligned}
F_k^{+}(\mathbf{u}_t; \boldsymbol{\gamma}, \boldsymbol{\Sigma}_t) &= \frac{1}{(2\pi)^{\frac{n}{2}} n} |\boldsymbol{\Sigma}_t|^{-\frac{1}{2}} \exp\left\{-\frac{\mathbf{u}_t' \boldsymbol{\Sigma}_t^{-1} \mathbf{u}_t}{2}\right\} \sum_{i=1}^{n} w_{ik} \left[1 + \sum_{s=1}^{m} \gamma_{is}^{k} \left(x_{it}^{s} - \mu_{is}^{+}\right)^{k}\right] && (21) \\
F_k^{*}(\mathbf{u}_t; \boldsymbol{\gamma}, \boldsymbol{\Sigma}_t, \nu) &= \frac{\Gamma\left(\frac{\nu+n}{2}\right) |\boldsymbol{\Sigma}_t|^{-\frac{1}{2}} \left[1 + \frac{\mathbf{u}_t' \boldsymbol{\Sigma}_t^{-1} \mathbf{u}_t}{(\nu-2)}\right]^{-\frac{\nu+n}{2}} \sum_{i=1}^{n} w_{ik} \left[1 + \sum_{s=1}^{m} \gamma_{is}^{k} \left(x_{it}^{s} - \mu_{is}^{*}\right)^{k}\right]}{n(\pi(\nu-2))^{\frac{n}{2}} \Gamma\left(\frac{\nu}{2}\right)} && (22) \\
\forall k &= 1, 2,
\end{aligned}$$

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 multivariate Normal and the MST are nested in (21) and (22), respectively, and that $F_k^{*}(\mathbf{u}_t; \boldsymbol{\gamma}, \boldsymbol{\Sigma}_t, \nu)$ tends to $F_k^{+}(\mathbf{u}_t; \boldsymbol{\gamma}, \boldsymbol{\Sigma}_t)$ as ν goes to infinity.

References

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

[2] Aslanidis, N. & Casas, I. (2013). Nonparametric Correlation Models for Portfolio Allocation. *Journal of Banking and Finance*, 37, 2268-2283.

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

[4] Bauwens, L., Laurent, S., and Rombouts, J. V. K. (2006). Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics*, 21, 79-109.

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

[6] Cappiello, L., Engle, R. F., and Sheppard, K. (2006). Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics*, 4, 537-572.

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

[8] Del Brio, E. B., Ñíguez, T. M., and Perote, J. (2011). Multivariate Semi-Nonparametric Distributions with Dynamic Conditional Correlations. *International Journal of Forecasting,* 27, 347-364.

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

[10] Engle, R. F., and Kelly, B. (2012). Dynamic Equicorrelation. *Journal of Business & Economic Statistics*, 30, 212-228.

[11] Engle, R. F., and Rangel, J. (2012). The Factor-Spline-GARCH Model for High and LowFrequency Correlations. *Journal of Business and Economic Statistics*, 30(1), 109–124.

[12] Fajardo, J., and Farias, A. (2010). Derivative Pricing Using Multivariate Affine Generalized Hyperbolic Distributions. *Journal of Banking and Finance*, 34, 1607-1617.

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

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

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

[16] Harris, R. D. F., and Mazibas, M. (2013). Dynamic Hedge Fund Portfolio Construction: A Semi-Parametric Approach. *Journal of Banking and Finance*, 37, 139-144.

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

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

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

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

[21] Ñíguez, T. M., and Perote, J. (2012). Forecasting Heavy-Tailed Densities with Positive Edgeworth and Gram-Charlier Expansions. *Oxford Bulletin of Economics and Statistics,* 74, 600-627.

[22] Ñíguez, T. M., and Perote, J. (2014). The Moment Expansions: A Semi-nonparametric Method with Applications for Risk Management. Westminster Business School Working Paper 14-1, 1-27.

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

[24] Patton, A. (2012). A Review of Copula Models for Economic Time Series. *Journal of Multivariate Analysis*, 110, 4-18.

[25] Polanski, A., and Stoja, E. (2011). Dynamic Density Forecasts for Multivariate Asset Returns. *Journal of Forecasting,* 30, 523-540.

[26] Sentana, E., Calzolari, G., and Fiorentini, G.. (2008). Indirect Estimation of Large Conditionally Heteroskedastic Factor Models, with an aApplication to the Dow 30 Stocks. *Journal of Econometrics*, 146, 10-25.

[27] Silvennoinen, A., and Terasvirta, T. (2009). Multivariate GARCH Models. In *Handbook of Financial Time Series*, ed. T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch, 201–229. New York: Springer.

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: $\sigma_{it}^2 = \alpha_{i0} + \alpha_{i1} u_{i,t-1}^2 + \alpha_{i2}\sigma_{i,t-1}^2,\ i = 1,2.$

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

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

	Π_1	Π_2	F_1^+	F_2^+	MST	F_1^*
			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)

	Π_1	Π_2	F_1^+	F_2^+	MST	F_1^*
			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	6090.16	6111.03

Notes: This table presents (Q)ML estimates of the parameters of the MME densities using Gaussian (F_k^+) and Student's t (F_k^*) as basis, $k = 2$ indicates that GNT transformations are implemented. The estimates of two versions of the MGC (Π_k) and the MST are also displayed, all distributions are bivariate. 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 weight 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.

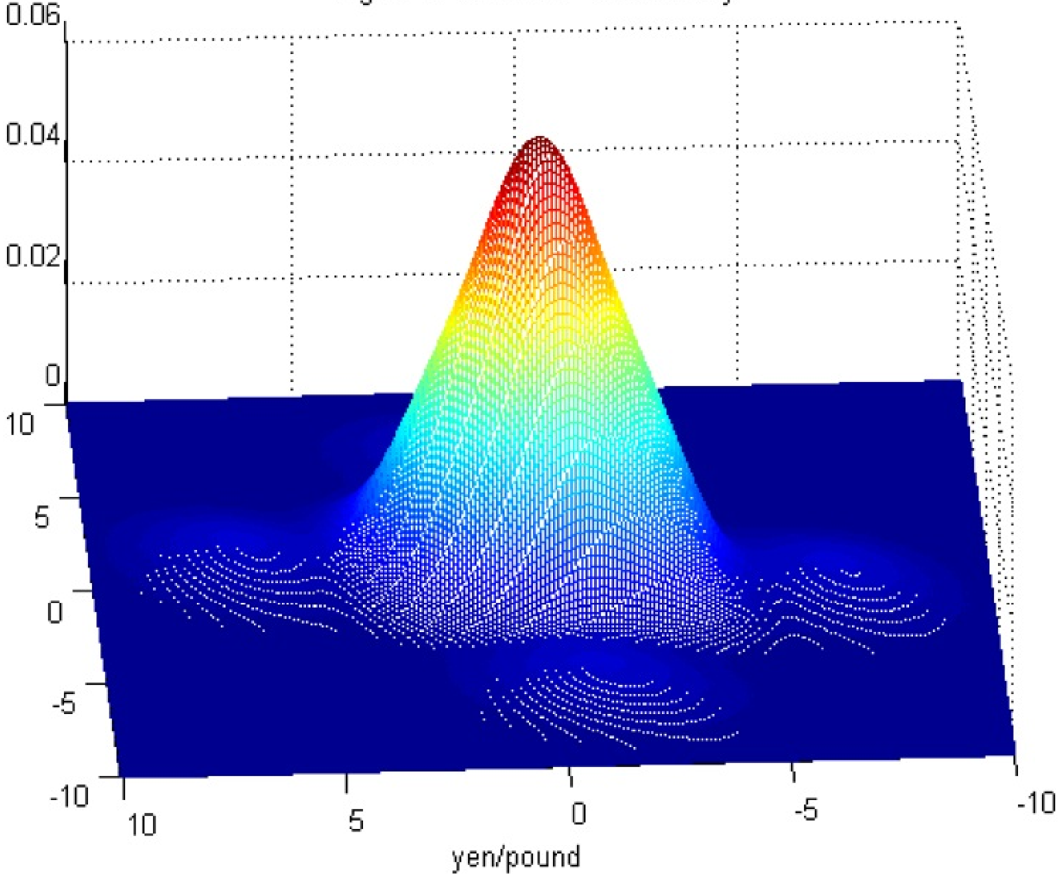

Figure 1: Bivariate ME density

Figure 2A: Marginal densities (yen/pound)

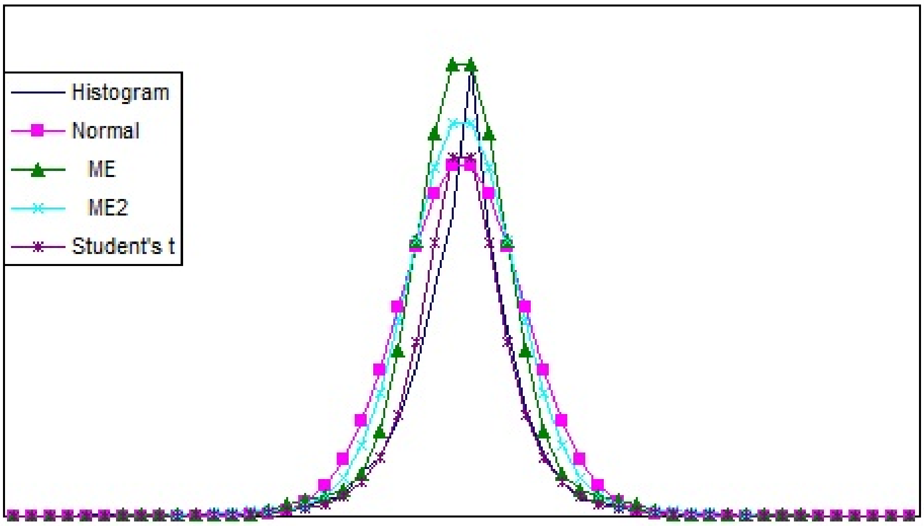

Figure 2B: Left talils of marginal densities (yen/pound)